THE PROCEEDINGS OF
THE PHYSICAL SOCIETY

Section B

VOL. 62, PART 8 1 August 1949 No. 356 B

The Effect of the Surrounding Atmosphere on the Delayed Fracture of Glass

BY C. GURNEY AND S. PEARSON
Royal Aircraft Establishment, Farnborough, Hants.

Communicated by G. Temple; MS. received 11th November 1948

ABSTRACT. Round soda-lime-silica glass rods were subjected to a series of given bending moments, and times to fracture were recorded. Experiments were made in air and *in vacuo* of 10^{-1} and 10^{-5} mm. of mercury; and the effects of heating *in vacuo* to drive off absorbed atmospheric constituents were also investigated. For the same prior heat treatment the slopes of the curves of stress against time to fracture decrease with decrease of pressure. At 10^{-5}mm. Hg, after heat treatment, the curve was very flat. It is evident that the main cause of delayed fracture is attack of the glass by atmospheric constituents. Experiments in which glass was subjected to air from which water vapour and carbon dioxide were removed separately and together showed that both these substances cause delayed fracture. Other constituents of the atmosphere seem relatively ineffective. These conclusions are in agreement with the work of Preston and his collaborators. It is further concluded that the atmospheric constituents can be supplied by capillary liquid contained in surface cracks and by migration from surface layers as well as by direct attack from the gaseous phase.

§ 1. INTRODUCTION

DELAYED fracture of glass has attracted much attention recently, including the notice of the Press (*The Times*, 12th February 1947). Among the more important papers giving accounts of experiments are those by Holland and Turner (1940) and Preston and his collaborators (Preston 1942, Baker and Preston 1946). Preston found that the average time for annealed soda-lime glass to fracture in bending under constant load increased from 10^{-2} to 10^4 seconds when the stress was reduced in the ratio 3 : 1. A previous paper by one of us (Gurney 1947) treated the theoretical aspects of the problem under the headings of approach to homogeneous and heterogeneous equilibrium. Homogeneous equilibrium is concerned with the arrangements of atoms in a given phase; as the strength of material depends on the atomic arrangement, gradual rearrangement of atoms in stressed glass will result in change of strength. A probable source of delayed fracture was considered to be atomic rearrangements in the highly stressed material at the ends of Griffith cracks. Heterogeneous equilibrium is concerned with the partition of material amongst more than one phase. Stress will alter the amount of moisture or other atmospheric constituents in equilibrium with the glass molecules and so lead to change of strength. Recent papers in which delayed fracture is attributed to atomic rearrangemnets in the glass are those by Murgatroyd (1944), who associated delayed fracture with the delayed elastic effect, and papers by Kontorova (1946), Saibel (1947), and Cox (1948). Poncelet (1945) and Taylor (1947) attribute delayed fracture both to atomic rearrangement in the glass and to attack by the atmosphere—the latter refers to the atmosphere as "catalysing" the atomic rearrangements. Smekal (1936), Preston (1942), and Orowan (1944) associate delayed fracture with atmospheric attack. To elucidate the problem further it is necessary to separate the two effects, and in the experiments now reported on, the effects of atmospheric attack were eliminated as far as possible by making delayed fracture experiments in a vacuum. Some experiments on these lines have already been reported on by Preston and Baker (1946). These investigators recorded whether or not fracture occurred in a given time interval. Their vacuum experiments were made at two durations of loading: 0·1 and 10 seconds. In our experiments a series of loads giving fractures in the time range 1–10^7 seconds was used.

§ 2. MATERIAL

The material tested was annealed soda–lime–silica glass rod of $\frac{1}{4}$ in. diameter having the following percentage analysis:

Silica	68·78	Alumina	2·76	Iron as ferric oxide	0·061
Lime	5·67	Titania	0·080	Magnesia	3·17
Boric oxide	0·43	Soda	17·95	Potash	0·48

It was received in lengths of 5 feet, the rods being laid in the packing case parallel to each other and in contact along their lengths. It had been stored in the laboratory in which it was tested for about a year before the experiments reported here were started. Photoelastic measurements indicated a residual internal stress of about 1,000 lb/in^2 with a standard deviation of 500 lb/in^2. More thorough annealing was not attempted as the same batch of rods was also used for cyclic fatigue tests reported elsewhere, in which it was most important that the rods should be straight. The presence of internal stress and of possible surface scratches due to packing and handling does not in any way invalidate the comparisons made in the paper, although the extra variability introduced by these factors necessitated more repetitions of each experiment than would otherwise have been necessary. It was considered that the advantage of readily available test pieces more than offset the work involved in carrying out a larger number of tests.

§ 3. APPARATUS

The glass rods were tested in four-point bending, the bending arm being 2 in. and the length subjected to uniform bending 3 in. For the vacuum experiments, the rods were placed in thin copper tubes whose walls had been rolled into a deep helix to reduce the bending stiffness, which was then small compared with that of the glass. It was found experimentally that 5% of the load was withstood by the copper tubes, and an allowance of this amount was made in calculating the loads to be applied. Each tube carried four internal brass bushes at the loading points, and load was applied through wire loops. Fracture was automatically recorded by the stopping of electric clocks. Vacuum was obtained by means of a mechanical backing pump and an oil diffusion pump, fitted with a P_2O_5 moisture trap. Preliminary heating was done in an electrically heated furnace.

In the first series of vacuum experiments, internally tinned copper U-tubes were placed in the line between the flexible copper tubes and the vacuum pump. By melting a piece of solder wire placed in the U-tubes, the tubes could be sealed. For subsequent vacuum experiments the flexible metal tubes containing the glass were coupled up six at a time to a 2½-in. diameter tube by means of a flanged coupling and polythene gasket. This tube was connected to the vacuum pump, which was kept running continuously during the experiment. In this series of tests the vacuum was continuously recorded by a Knudsen gauge which was placed at the end of a flexible metal tube, welded to the 2½-in. diameter tube at the end remote from the pump. The object of placing the gauge in this position was to ensure that the recorded pressure would be no lower than that in the tubes containing the glass. The main features of the Knudsen gauge were according to the design of Dumond and Pickels (1935).

§ 4. EXPERIMENTS

Four hundred test pieces were numbered consecutively and allocated to the various tests, according to tables of random numbers. A series of delayed fracture experiments was made in air at room temperature. Twelve tests were made at each stress, and the median times to fracture (circles) are plotted in the lower curve of Figure 1. The effect of heat treatment *in vacuo* on the strength when subsequently exposed to the atmosphere was investigated. The median points for glass so treated are shown by crosses on the lower part of Figure 1. They do not depart significantly from the curve drawn through previously obtained points, nor did the mean strength of quick-loading tests on twenty-four glass rods, twelve of which had previously had the vacuum heat treatment, show a significant difference.

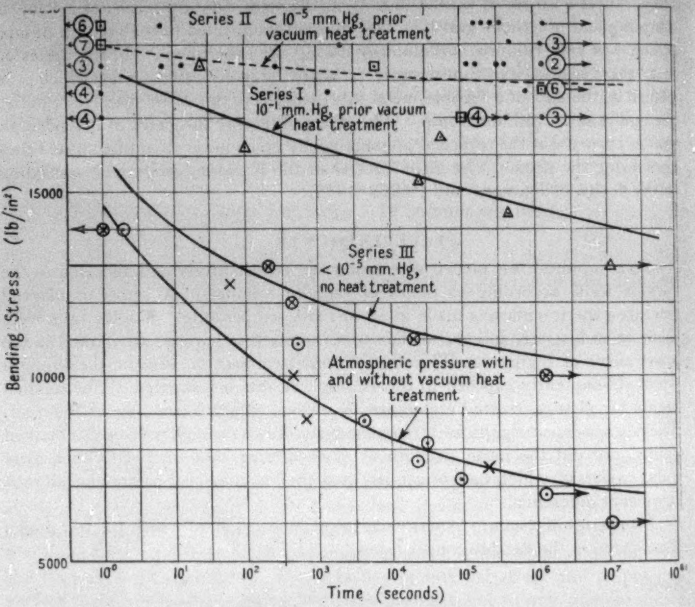
Five series of vacuum experiments were made, in all of which the test pieces were tested in the flexible copper tubes.

Series I

For this series of tests, copper U-tubes, each containing a length of solder wire, were placed in the pipe line between the diffusion pump and the flexible tubing. The glass was heated for 3 hours at 310 ± 10° c. in a vacuum of 10^{-4} mm. mercury, as measured by a Macleod gauge fitted with a cold trap. While the vacuum pumps were still running, the U-tubes were heated so as to melt the solder wire, which ran down to the bottom of the U and sealed the tubes. The sealed tubes were then allowed to cool and were used in the delayed fracture experiments. By mounting a vacuum gauge on the end of the flexible tube furthest from the pump the vacuum remaining in the tubes after sealing could be measured. This proved to be of the order of 10^{-1} mm. Hg. The deterioration from 10^{-4} mm. Hg before sealing to 10^{-1} after sealing was presumably due to gassing from the hot metal. The results of the delayed fracture experiments are shown in Figure 1. Twelve tests were made at each stress and the median times to fracture are represented by triangles.

Series II

Continuous evacuation was used for this and for subsequent series of tests. For this series the test pieces were first heated to 360 ± 10° c. for 48 hours. After one hour's heating the vacuum recorded by the Knudsen gauge at the end of the pipe line was about 5×10^{-5} mm. Hg, which reduced to less than 1×10^{-5} mm. Hg after 48 hours' heating. The glass was allowed to cool and then the furnaces were removed and the glass was tested, the pressure being maintained below 1×10^{-5} mm. Hg by continuous evacuation. The delayed fracture results are represented at the top of Figure 1. As few delayed fractures occurred, all the results are shown, the median points being represented by squares. It will be seen that most test pieces either failed during the application of load or remained unbroken. The standard deviation of the time to fracture on a log scale is of the order of 5·0 at a vacuum of 10^{-5} mm. Hg and 0·5 in air at atmospheric pressure. These figures apply to mean times to fracture of about the middle of the range shown



⊙ Median points for tests at atmospheric pressure with no prior vacuum treatment. (12 tests at each stress.)

× Median points for tests at atmospheric pressure after prior vacuum heat treatment. (12 tests.)

⊗ Median points for tests in vacuum of $< 10^{-5}$ mm. Hg after prior vacuum treatment at room temperature (6 tests).

△ Median points for tests in vacuum of 10^{-1} mm. Hg after prior vacuum heat treatment (12 tests).

⊡ Median points for tests in vacuum of $< 10^{-5}$ mm. Hg after prior vacuum heat treatment (12 tests).

• Results of individual tests.

←Ⓐ "A" denotes number of test pieces breaking in less time.

Ⓐ→ "A" denotes number of test pieces which did not break.

Figure 1.

in the Figures. We have made the plausible assumption that the slopes of the curves are proportional to standard deviations, and have accordingly drawn the dotted curve at the top of Figure 1 to have a mean slope of about 0·1 times the mean slope of the bottom curve.

Series III

This series of experiments was designed to estimate the importance of heating the glass *in vacuo* on the delayed fracture effect. The test pieces were maintained in a vacuum of less than 10^{-5} mm. Hg for 48 hours without any heating, after which loads were applied and the times to fracture in the vacuum were recorded. The crossed circles in Figure 1 represent median points for six nominally identical test pieces.

Series IV

In order to investigate what constituents of the atmosphere were effective in causing delayed fracture, six test pieces were heated *in vacuo* for 48 hours as in the second series of vacuum experiments, and then, when the test pieces were cool, air dried by passing through P_2O_5 tubes was admitted to the system. Loads corresponding to a stress of 18,400 lb/in^2 were applied, the time from first admitting the dried air to applying the load being about 60 minutes. If the vacuum had been maintained, the median time to fracture would have been of the order of 10^4 seconds at this stress, but in the presence of the dry air all test pieces failed before the full load was applied. When air from which CO_2 had been extracted by passing through caustic potash followed by Baryta water was admitted after the vacuum heat treatment, all test pieces again failed without withstanding the full load; but when both water vapour and CO_2 were extracted by passing the air through caustic potash followed by P_2O_5, the median time to fracture at atmospheric pressure was of the same order as that in a vacuum of less than 10^{-5} mm. Hg.

Series V

In this series of experiments the effect on the short-time loading strength of maintaining the glass *in vacuo* at room temperature was investigated. Unloaded glass rods were maintained in a vacuum of less than 10^{-5} mm. Hg at room temperature for periods corresponding to the abscissae of Figure 2 and then fractured at a constant rate of loading of about 20,000 lb/in^2/min. The points shown are median points for twelve repetitions of each experiment. For comparison twelve tests were made in which the glass was heated to $360 \pm 10°$ C. in a vacuum of less than 10^{-5} mm. Hg for 48 hours as in the second series of vacuum experiments. The test pieces were then allowed to cool, the furnaces were removed, and the rods were fractured at about 20,000 lb/in^2/min. The median point for these tests is also shown in Figure 2. The glass used for this series of experiments was specified to the same batch formula as that used previously, but was not from the same batch.

§ 5. DISCUSSION

In accordance with a previous paper (Gurney 1947), delayed fracture is considered to be due to the gradual spread of cracks. It is immediately clear from Figure 1 that the main cause is associated in some way with the presence of the atmosphere. At a vacuum of less than 10^{-5} mm. Hg, and after heating the glass, only one median point lay in the range 10–10^6 seconds. It is evident that the curve of stress–time has become very flat. Atomic rearrangements made possible by stress only and unaided by atmospheric constituents do not therefore lead to an appreciable delayed fracture effect. The experiments in air from which water vapour and carbon dioxide were excluded in turn indicate that both these substances cause appreciable delayed fracture effects. The other constituents of the atmosphere seem relatively ineffective.

A plausible extrapolation of the curves to the left suggests that they might intersect at a stress of about 20,000 lb/in². This stress could be interpreted as the strength of glass when free from atmospheric attack, either because no atmosphere was present or because the time was too short for appreciable attack to take place. Baker and Preston (1946) found that the strength *in vacuo* was about the same as in air at a 10^{-2} second loading period. It should be noted that the rate of attack is not proportional to the pressure, as might be expected from too simple reasoning. If it were, then the curves of Figure 1 would be expected to be identical in shape but bodily displaced relative to each other, and having a common asymptote. The fact that the curves do not tend to a common asymptote suggests that at the longer times an effective equilibrium is reached between glass and atmospheric constituents, the equilibrium composition varying with stress and pressure.

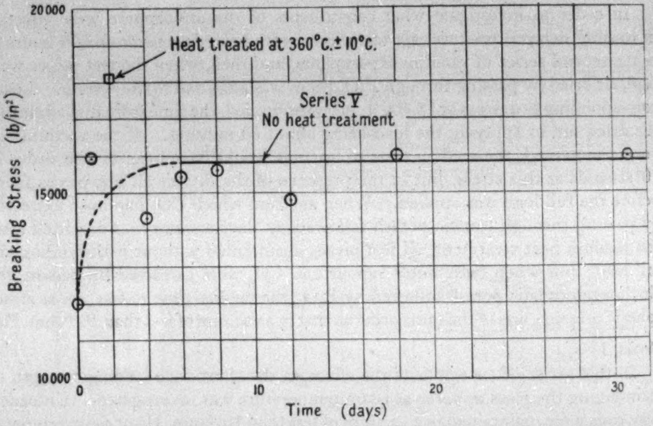


Figure 2. Pressure $< 10^{-5}$ mm. Hg. Breaking stress when loaded to failure at 20,000 lb/in²/min. Each point shown is the mean of 12 tests.

The water and carbon dioxide which cause the attacks at the ends of the cracks may be derived from any of three phases: from the gaseous phase, from capillary liquid held in surface cracks, or from the complex layer of glass and atmospheric constituents which forms the surface of normal glass. The formula relating equilibrium vapour pressure P to radius of meniscus (assumed cylindrical) of the capillary liquid r is

$$\log(P_r/P_\infty) = V_L\gamma/\mathbf{R}Tr,$$

where V_L is specific volume of liquid and γ is surface tension. The smallest possible value for r is of the order of atomic dimensions (10^{-8} cm.) For this value, capillary water would not be present at partial pressures of water less than 0·1 mm. Hg. In ordinary atmosphere, attack from capillary water may be important, but it is not likely to be of consequence in the vacuum experiments.

The fact that delayed fracture occurred in a vacuum of 10^{-5} mm. Hg when the glass had not been heated, but did not occur to an appreciable extent after the glass had been heated, shows that direct attack by the gaseous phase is not the only factor in the vacuum experiments. The evidence of Figure 2 is also relevant in that the short-time loading strength after a month in the high vacuum without heating did not reach the short-time loading strength of glass heated for 48 hours *in vacuo*. It is evident that some of the delayed fracture is due to atmospheric constituents contained in the surface layers of the glass. Presumably, when the glass is stressed, atoms of atmospheric constituents diffuse along the surface layers to the ends of the cracks. It should be noted that heating *in vacuo* does not cause a permanent change in the strength of glass re-exposed to the atmosphere. This is shown by the crosses on the lower curve in Figure 1, which are scattered about the curve for glass which has not been heated.

Some idea of the relative importance of direct attack by the atmosphere and of attack by capillary liquid and diffusion in surface layers can be obtained from the heat of reaction of the process. From Figure 1, and at the longer periods of loading, we may make a rough estimate of $(\partial X/\partial P)_T$ at atmospheric pressure, where X is stress.

Vonnegut and Glathart (1946) have plotted the breaking strength of glass of similar composition against temperature for three different periods of loading. From their data it is possible to make a rough estimate of $(\partial X/\partial T)_P$ at atmospheric pressure and room temperature and long-period loading. By making use of the identity

$$\frac{dP}{dT} = -\frac{\partial X/\partial T}{\partial X/\partial P}$$

an estimate of $(\partial P/\partial T)_X$ is obtained. If we regard the reaction as glass + atmosphere = complex of glass and atmosphere, i.e. g + a = ga, then $\mu_g + \mu_a = \mu_{ga}$ where μ is chemical potential.

Differentiating with respect to T and P, we obtain $dP/dT = L/T\Delta\overline{V}$, where L is the heat of the reaction and $\Delta\overline{V}$ is the difference between the partial molar volumes before and after the reaction. If we regard the atmosphere as a perfect gas and neglect the volumes of the condensed phases, $\Delta\overline{V} = \boldsymbol{R}T/P$. We thus obtain

$$L = \frac{dP/dT}{P}\boldsymbol{R}T^2.$$

Making an estimate of dP/dT as described above, we obtain $L \simeq -700$ cal/mole. This value for L is extremely approximate, and should be interpreted merely as showing that the heat of reaction is small and negative, i.e. heat must be supplied to maintain the temperature constant. Heats of adsorption from gaseous phases are always positive, due to the large entropy of gaseous molecules. At room temperature and atmospheric pressure the reaction is not, therefore, one between glass and a gaseous phase. It must be between the glass and condensed atmospheric constituents contained in capillary liquid and surface layers.

At higher temperatures (above 250° c.), Vonnegut and Glathart's results show a change in sign of $(\partial X/\partial T)_P$. If, as seems reasonable, the sign of $(\partial X/\partial P)_T$ does not show a corresponding change of sign, then the heat of reaction would become positive at temperatures above 250° c. This would be in accordance with the expectation that at elevated temperatures capillary water and water in surface layers has been driven off and that reaction is taking place directly with the gaseous phase. Delayed fracture appears, therefore, to be caused by approach both to homogeneous and heterogeneous equilibrium; but the homogeneous equilibrium entails atomic migrations in a surface layer of glass and atmospheric constituents, and not simply in the glass itself.

REFERENCES

Baker, T. C., and Preston, F. W., 1946, *J. Appl. Phys.*, **17**, 179.
Cox, S. M., 1948, *J. Soc. Glass Tech.*, **32**, 127.
Dumond, J. W. M., and Pickels, W. M., 1935, *Rev. Sci. Instrum.*, **6**, 362.
Gurney, C., 1947, *Proc. Phys. Soc.*, **59**, 169.
Holland, A. J., and Turner, W. E. S., 1940, *J. Soc. Glass Tech.*, **24**, 46.
Kontorova, T. A., 1946, *C.R. Acad. Sci., U.S.S.R.* (*Doklady*), **54**, 1.
Murgatroyd, J. B., 1944, *J. Soc. Glass Tech.*, **28**, 406.
Orowan, E., 1944, *Nature, Lond.*, **154**, 341.
Poncelet, E. F., 1945, *Coll. Chem.*, **6**, 77.
Preston, F. W., 1942, *J. Appl. Phys.*, **13**, 623.
Saibel, E., 1947, *J. Chem. Phys.*, **15**, 760.
Smekal, A., 1936, *J. Soc. Glass Tech.*, **20**, 432.
Taylor, N. W., 1947, *J. Appl. Phys.*, **18** (11), 943.
Vonnegut, B., and Glathart, J. L., 1946, *J. Appl. Phys.*, **17**, 1082.

The Electrical Properties of Bismuth Oxide

By R. MANSFIELD
Bedford College for Women, University of London

Communicated by H. T. Flint; MS. received 29th October 1948

ABSTRACT. The conductivity σ and thermo-electric power dE/dT of bismuth oxide is measured over a temperature range of 680° to 150° c., and for variations of oxygen pressures from 76 to 10^{-4} cm. Hg. The results above 340°c. and at atmospheric pressure are well represented by the formulae $\sigma=A_1\exp(-\alpha/2T)$ and below 340 ° c. by $\sigma=A_2\exp(-\alpha/T)$, in which the values of the constants A_1, A_2 and α are very similar for two different specimens. The variation of σ with oxygen pressure above 500° c. and for pressures down to 1 mm. Hg obeys the formula $\sigma=k(P_{O_2})^{\frac{1}{4}}$. The thermo-electric power results may be summarized by the equation $dE/dT=a+b/T$, where a and b are constants, the positive sign of dE/dT indicating that Bi_2O_3 is a defect semiconductor. The variations of dE/dT with oxygen pressure are small less than 10%, except at pressures less than 10^{-3} cm. Hg, when the sign of dE/dT changes at temperatures greater than 550° c. Under these conditions it is suggested that Bi_2O_3 becomes an excess or intrinsic semiconductor. The results are discussed and good agreement is obtained between theory and experiment if it is assumed that normal Wagner and Schottky lattice defects become of importance at low temperatures when the concentration of free electrons is small.

§1. INTRODUCTION

The theory of semiconductors as developed on the basis of the simple model proposed by A. H. Wilson in 1931 explains in general terms the qualitative properties of a considerable number of semiconductors. Agreement between theory and experiment is, however, only of the right order of magnitude. For example, with zinc oxide, Seitz (1940) has calculated a value of the thermoelectric effect, using measurements on the conductivity of zinc oxide, but the predicted value agrees only very roughly with the observed values. The reasons for the discrepancies are as follows. Firstly, the properties of semiconductors are largely dependent on the concentration of impurity present in a given specimen, and consequently consistency between theory and experiment of the different electrical properties of the same substance can only be expected if all measurements are made on the same specimen, a difficulty which can be eliminated by suitably designed experiments. Secondly, the theory assumes a model in which the number of free electrons in the material is sufficiently small to allow the methods of classical statistics to be used in calculating the properties of the electrons. Shifrin (1944) has discussed and developed the theory of the modifications which have to be made if this condition is not satisfied. The purpose of the present investigation is to produce evidence that good agreement between theory and experiment can be obtained for a highly resistive semiconductor, such as bismuth oxide, in which the number of free electrons is small, so that the theory based on Wilson's model can be used. A difficulty arises in interpreting the results of Bi_2O_3 because the concentation of free electrons becomes so small that ordinary lattice defects frozen into the lattice play a part in the process of conduction. The required modification to the theory has been developed by Nijboer (1939).

§2. EXPERIMENTAL PROCEDURE

A sample of pure bismuth oxide (BPC) was pressed into the shape of a cylinder of diameter $\frac{3}{4}$ in. and length $\frac{3}{4}$ in. using a steel mould and a pressure of about 200 kg/cm². This was then sintered at 760° c. for 14 hours. The resultant specimen had a density of 8·7 gm/cm³, approximately 95% that of the pure crystal.

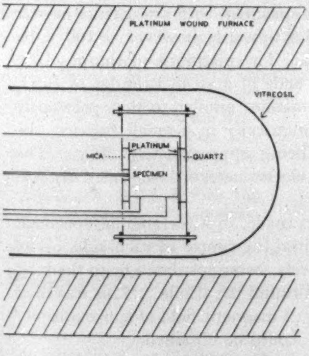


Figure 1.

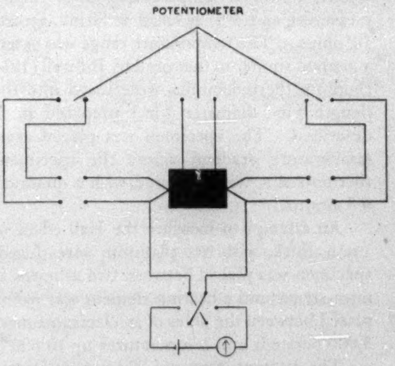


Figure 2.

After grinding the end faces flat, two platinum wires were fused into the side of the specimen, using a fine jet of an oxy-coal-gas flame. These wires were used as potential probes in measurements of the conductivity. The specimen was then placed between two platinum discs to which were attached chromel-alumel thermojunctions, and the whole was pressed tightly together in a holder constructed of nickel. Quartz discs were used to insulate the nickel holder from the specimen (Figures 1 and 2).

The conductivity was measured by passing a known current through the specimen and measuring the potential difference between the probes. At high temperatures the potential difference was measured using a potentiometer, and at low temperatures, when the resistance of the specimen was greater than 10^5 ohms, and a potential difference of 120 volts was applied across the specimen, a quadrant electrometer was used with the needle connected to one of the quadrants. The procedure consisted in measuring the potential of either probe with respect to one end of the specimen, the difference being the potential required. Readings were taken with the current flowing through the specimen first in one direction and then in the other. The temperature of both ends of the specimen was measured by means of chromel-alumel junctions before and after making the conductivity measurements.

For thermo-electric measurements a temperature gradient of approximately 50° c. was obtained by placing the specimen eccentrically in the furnace. Over the range 250° to 720° c. the thermo-E.M.F. was satisfactorily measurable by means of a potentiometer.

The effect of oxygen pressure on both the conductivity and the thermo-electric power was measured over a wide range of pressures, 76 cm. to 10^{-4} cm. Hg. Care was taken that the specimen had reached equilibrium with the surrounding atmosphere before readings were taken, and in most cases this involved leaving the system for 14 hours at one pressure. The electrical properties of the specimen remained the same after several days, in which the specimen was kept at 650° c. at low pressures. A complete set of readings was taken for two specimens prepared from two different samples of bismuth oxide.

With this simple method it was found impossible to measure the thermo-electric power at temperatures lower than 250° c., because of the difficulty of measuring an E.M.F. as small as 50 mv. across a specimen with a resistance of over 10^5 ohms. The temperature range was extended upwards, however, to 100° c. by a method similar to that used by Bidwell (1914). Two platinum–platinum + 13% rhodium thermocouples were fused into the ends of a small cylinder of Bi_2O_3 (length $\frac{3}{8}$ in., diameter $\frac{3}{16}$ in.) prepared in a manner similar to that previously described. The specimen was placed asymmetrically in a small furnace, the temperature gradient across the specimen being approximately 30° c. The thermo-E.M.F. was measured with a quadrant electrometer whose sensitivity was 0·3 mm/mv.

An attempt to measure the Hall effect was made, in a rectangular specimen 1 mm. thick, with two platinum wires fused into the centre of each side. The specimen was placed between two asbestos sheets, being insulated from them by mica-strips, and a heating element was wound round the whole. This was then placed between the poles of an electromagnet. A magnetic field of approximately 5,000 oersteds and temperatures up to 600° c. could be obtained.

The attempt, however, did not succeed, since the system was not sufficiently stable to measure an E.M.F. of less than 5 mv. across the specimen at a temperature of 500° c. As will be shown later, the theoretical value at this temperature is of the order 0·6 mv.

§ 3. RESULTS

There was close agreement between the values of the conductivity obtained from the two specimens. The variation of conductivity σ with temperature T can be summarized by equations

$$\sigma = A_1 \exp(-\alpha/2T), \qquad T > T_c \qquad \ldots\ldots(1)$$

$$= A_2 \exp(-\alpha/T), \qquad T < T_c, \qquad \ldots\ldots(2)$$

where A_1, A_2 and α are constants and T_c is approximately 340° c. Figure 3 illustrates the variation of conductivity with temperature at different oxygen pressures.

The two methods of measuring the thermo-electric power gave values agreeing to within 10%; the most likely explanation of the difference is error in measuring the actual temperature of the contacts. In the first method the temperature difference between the platinum discs was probably slightly greater than that across the specimen. The variation of thermo-electric power dE/dT with temperature obeys the empirical relation

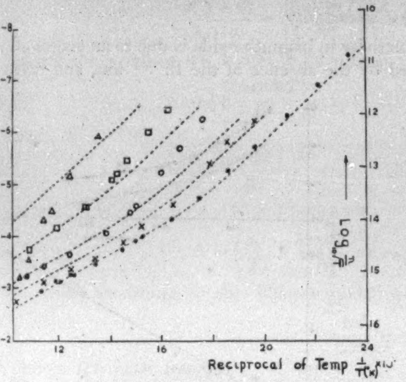


Figure 3. Variation of conductivity with temperature at different oxygen pressures.

• • • •	76 cm. Hg.
× × × ×	7·7 cm. Hg.
○○○○	1·22 cm. Hg.
□□□□	1·48 × 10^{-2} cm. Hg.
△ △ △ △	10^{-4} cm. Hg.
Broken curves	Theoretical.

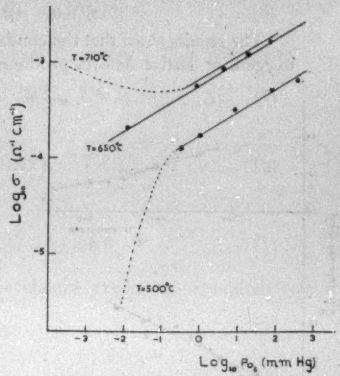


Figure 4. Variation of conductivity with oxygen pressure at fixed temperatures.

$$\frac{dE}{dT} = a + \frac{b}{T}, \qquad \ldots\ldots(3)$$

where a and b are constants (see Figures 5, 6).

The sign of the thermo-electric power was positive, suggesting conduction by holes, a conclusion which was corroborated when measurements of the variation of conductivity with oxygen pressure were taken. It was then found that the conductivity decreased with decreasing oxygen pressure. Thus bismuth oxide is a defect or abnormal type of semiconductor.

The variation of conductivity with oxygen pressure obeys the relationship

$$\sigma = C(P_{O_2})^{1/n}, \qquad \ldots\ldots(4)$$

where $n = 4$ and C is a constant, over the range of temperatures where equation (1) holds, and down to pressures of 1 mm. Hg. The variation of thermo-electric power with oxygen pressures was less than 10%, except for pressures less than

10^{-3} cm. Hg. At these low pressures dE/dT changes sign as the temperature is increased above 550° c. Under these conditions bismuth oxide appears to become an excess, or an intrinsic semiconductor.

Measurements were made to check Ohm's law at high temperatures. The relation between the current flowing through the specimen and the potential difference between the probes was found to be linear from 1·6 volts to 0·013 volt.

§ 4. DISCUSSION OF RESULTS

(i) *Conductivity*

The results show that the conductivity in bismuth oxide is due to an excess of oxygen, or lattice defects produced by the absence of the Bi^{+++} ion, and that conduction is by positive holes. The action of the deviations from stoichiometric proportion in producing free holes may be represented by the equation

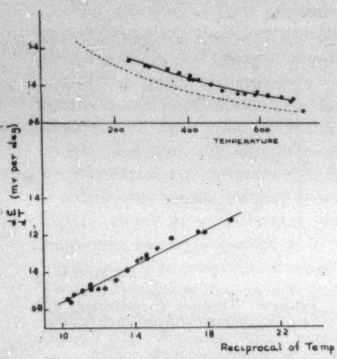


Figure 5. Variation of the thermo-electric power with temperature for the same specimen of bismuth oxide as used for conductivity measurements.

Full curve, experimental ; broken curve, theoretical.

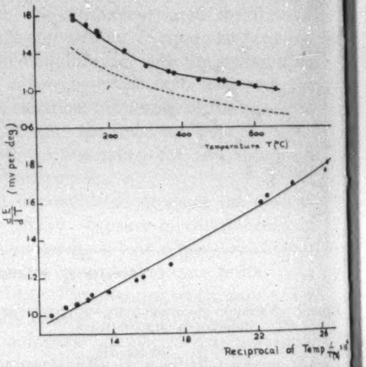


Figure 6. Variation of thermo-electric power wi temperature using thermocouples fused into t specimen.

Full curve, experimental ; broken curve, theoretic

$$\text{free hole} + \text{negative ion} \rightleftharpoons \text{bound hole.}$$

If n = number of free holes, n_0 = number of impurity atoms, and n' = number of negative ions with which the free holes can combine, then the law of mass action gives

$$\frac{nn'}{n_0 - n} = K \exp(-\Delta E/\boldsymbol{k}T), \qquad \ldots\ldots(5)$$

where ΔE = energy of activation and K is a constant.

Now according to Wagner and Schottky (1930), for an ionic crystal in thermal equilibrium there are a number of vacant lattice sites of both constituents, their number being given by $N = N_0 \exp(-W/\boldsymbol{k}T)$, where N_0 = number of ions of a given kind per cm^3 and

$$W = \tfrac{1}{2}(W_+ + W_- - W_L),$$

where W_+ and W_- are respectively the energies required to remove a positive and negative ion out of the crystal, and W_L is the lattice energy per ion pair. During the process of sintering as the specimen cools a number of these defects will be frozen in.

The number n' of negative ions with which a free hole can combine will consequently be $N+n$, and thus (5) becomes

$$\frac{n(N+n)}{n_0-n} \sim \exp(-\Delta E/\boldsymbol{k}T). \qquad \ldots\ldots(7)$$

Nijboer (1939) has shown, by considering the free energy of the system, that

$$\frac{n(N+n)}{n_0-n} = \left\{\frac{2\pi \boldsymbol{mk}T}{\boldsymbol{h}^2}\right\}^{3/2} \exp(-\Delta E/\boldsymbol{k}T), \qquad \ldots\ldots(8)$$

where the symbols $\boldsymbol{m}$, $\boldsymbol{k}$ and $\boldsymbol{h}$ have their usual meanings. This reduces to

$$n = n_0^{\frac{1}{2}}\left\{\frac{2\pi \boldsymbol{mk}T}{\boldsymbol{h}^2}\right\}^{\frac{3}{4}} \exp(-\Delta E/2\boldsymbol{k}T), \quad T \gg T_c, \qquad \ldots\ldots(9)$$

and

$$n = \frac{n_0}{N}\left\{\frac{2\pi \boldsymbol{mk}T}{\boldsymbol{h}^2}\right\}^{3/2} \exp(-\Delta E/\boldsymbol{k}T), \quad T \ll T_c, \qquad \ldots\ldots(10)$$

where T_c is given by $n=N$ and hence by

$$N = \left(\frac{n_0}{2}\right)^{\frac{1}{2}}\left\{\frac{2\pi \boldsymbol{mk}T}{\boldsymbol{h}^2}\right\}^{\frac{3}{4}} \exp(-\Delta E/2\boldsymbol{k}T_c). \qquad \ldots\ldots(11)$$

Now according to the Wilson model we should expect the conductivity

$$\sigma = \frac{4nl_0\boldsymbol{e}}{3\sqrt{(2\pi \boldsymbol{mk}T)}}, \qquad \ldots\ldots(12)$$

where l_0 = mean free path.

When we combine (12), (9) and (10) we obtain equations similar to (1) and (2).

If we choose $N=3\cdot16\times10^{15}$, $\Delta E=0\cdot9$ ev. and $n_0=10^{19}/\mathrm{cm}^3$ very close agreement is obtained between the variation with temperature of $\log\sigma$, at atmospheric pressure, and $\log(n/\sqrt{T})$ (see Figure 3). Since Figure 4 shows that, over the range of temperature above T_c, $\log\sigma \simeq (P_{O_2})^{\frac{1}{4}}$ for a given temperature, where P_{O_2} is the oxygen pressure, combining equations (12) and (9) we should expect

$$n_0 \propto (P_{O_2})^{\frac{1}{4}}. \qquad \ldots\ldots(13)$$

The Table gives values of n deduced from equation (8) for the different oxygen pressures at which measurements were made, assuming $n_0=KP_{O_2}{}^{\frac{1}{4}}$, and $n_0=10^{19}$ at atmospheric pressure. Figure 3 illustrates the variation of $\log(n/\sqrt{T})$ with $1/T$; this compares favourably with the experimental points of Figure 3.

Number of Free Holes/cm³ for Various Values of n_0
($N=3\cdot16\times10^{15}/\mathrm{cm}^3$)

T(°c.) \ n_0	10^{19}	$3\cdot16\times10^{18}$	$1\cdot3\times10^{18}$	$1\cdot4\times10^{17}$	10^{16}
727	$1\cdot5\times10^{17}$	$8\cdot5\times10^{16}$	$5\cdot24\times10^{16}$	$1\cdot76\times10^{16}$	$3\cdot4\times10^{15}$
527	$3\cdot28\times10^{16}$	$1\cdot77\times10^{16}$	$1\cdot07\times10^{16}$	$2\cdot7\times10^{15}$	$3\cdot3\times10^{14}$
327	$1\cdot29\times10^{15}$	$7\cdot9\times10^{14}$	$3\cdot5\times10^{14}$	$4\cdot3\times10^{13}$	$3\cdot16\times10^{12}$
394	$6\cdot3\times10^{15}$	$3\cdot16\times10^{15}$	$1\cdot6\times10^{15}$	$2\cdot7\times10^{14}$	$2\cdot1\times10^{13}$
283	$5\cdot85\times10^{14}$	$2\cdot13\times10^{14}$	$8\cdot8\times10^{13}$	$9\cdot5\times10^{12}$	$7\cdot0\times10^{11}$
227	$7\cdot6\times10^{13}$	$2\cdot4\times10^{13}$	$9\cdot6\times10^{12}$	$1\cdot03\times10^{12}$	$7\cdot6\times10^{10}$
127	$3\cdot16\times10^{11}$	10^{11}	$4\cdot0\times10^{10}$	$4\cdot3\times10^{9}$	$3\cdot16\times10^{8}$

(ii) *Thermo-electric Effect*

According to Shifrin (1944), for an ionic lattice the thermo-electric power at high temperatures for a semiconductor is given by

$$\frac{dE}{dT}=\frac{\boldsymbol{k}}{\boldsymbol{e}}\left\{-\frac{\mu}{\boldsymbol{k}T}+\frac{3}{2}\frac{\int_0^\infty \chi^2 f_0\,d\chi}{\int_0^\infty \chi f_0\,d\chi}\right\}, \qquad \ldots\ldots(14)$$

in which μ is the chemical potential of the electrons and f_0 is the Fermi-Dirac function. This can be simplified by using the classical value of $f_0=e^{-\chi}$; then

$$\frac{\int_0^\infty \chi^2 e^{-\chi}\,d\chi}{\int_0^\infty \chi e^{-\chi}\,d\chi}=2 \qquad \ldots\ldots(15)$$

and

$$\frac{dE}{dT}=\frac{\boldsymbol{k}}{\boldsymbol{e}}\left\{-\frac{\mu}{\boldsymbol{k}T}+3\right\}. \qquad \ldots\ldots(16)$$

Seitz (1940) has shown that for $\Delta E \gg \boldsymbol{k}T$, μ becomes very nearly equal to $-\Delta E/2$. The theoretical curves of Figures 5 and 6 were obtained by substituting for μ the value of ΔE obtained by conductivity measurements. Conversely, if dE/dT is plotted against $1/T$, the value of μ obtained is 0·488 ev. as compared to $\Delta E/2=0{\cdot}45$ ev. obtained by conductivity measurements. Also, as dE/dT was found to be nearly independent of oxygen pressure except at low pressures and high temperatures, equation (16) indicates that μ does not vary appreciably with the concentration of impurity centres. This is to be expected if μ is nearly equal to $-\Delta E/2$.

(iii) *Hall Effect*

Using the above values of n, we see that at 500° c. the Hall coefficient R is given by

$$R=-\frac{3\pi}{8}\frac{1}{\boldsymbol{nec}}.$$

Remembering that conduction is by holes, and converting to practical units, at 500° c. we should expect

$$R=380\times 10^{-8}\ \text{volt cm/amp. gauss.}$$

Using a field of 4,000 oersted and a current of 4 ma., for the specimen used we should expect an E.M.F. of 0·6 mv. across the specimen. This was not measurable with the apparatus used.

§ 5. CONCLUSION

For a highly resistive semiconductor such as bismuth oxide, good agreement is obtained between theory and experiment. The relationship between conductivity and temperature can be explained if it is assumed that the normal Wagner and Schottky lattice defects become of importance when the concentration of free electrons becomes small.

ACKNOWLEDGMENTS

The author wishes to express his thanks to Professor J. P. Andrews for suggesting this work on highly resistive semiconductors, and for encouragement and advice at all times.

REFERENCES

Bidwell, C. C., 1914, *Phys. Rev.*, **3**, 204.
Nijboer, B. R. A., 1939, *Proc. Phys. Soc.*, **51**, 575.
Seitz, F., 1940, *Modern Theory of Solids* (New York and London: McGraw-Hill Book Co., Inc.).
Shifrin, K., 1944, *J. Phys. U.S.S.R.*, **8**, 242.
Wagner, C., and Schottky, 1930, *Z. Phys. Chem.* B, **11**, 163.

The Mechanism of Dilatancy

By E. N. da C. ANDRADE and J. W. FOX
University College, London

MS. received 26*th March* 1949; *read* 4*th February* 1949

ABSTRACT. The dilatancy in question is the change of overall volume produced by strain in an assembly of particles. This has been demonstrated in the case of a mixture of sand and water by a photoelectric method of measuring the wetness of the surface. The general phenomenon has been experimentally studied in detail with a two-dimensional hexagonal array of uniform cylinders, part of the free surface of which is loaded with a rigid piston. An alternate increase and decrease of overall volume as the piston descends has been demonstrated and the mechanism of slip by which it comes about has been established. The elastic stress distribution in the array before slipping takes place has been calculated and shown to be a determining factor. The effect of friction between the component units has been studied. Experiments have also been carried out on three-dimensional irregular arrays of carbon shot and of sand.

§1. INTRODUCTION

The term "dilatancy" is used today in two senses. On the one hand Freundlich and Röder (1938) have applied it to characterize the behaviour of certain suspensions which show an increase of viscosity with increasing rate of shear. On the other hand, the word is still used as it was originally by Osborne Reynolds (1885), to denote a particular type of behaviour exhibited by a collection of particles in contact with one another. In his own words, "the property consists of a definite change of bulk, consequent on a definite change of shape or distortional strain, any disturbance whatever causing a change of volume and generally dilatation". It is noteworthy that Osborne Reynolds was less interested in the phenomenon as a property of ordinary gross matter than for its possible bearing on the properties of the ether of space, which at that time dominated cosmic theory. "The discovery, if I may so call it, was the result of an attempt to conceive the mechanical properties a medium must possess in order that it might fulfil the functions of an all-pervading ether" (1886). Similarly, Newton's investigation into the properties of a fluid endowed with frictional properties, which has given rise to the term "Newtonian viscosity", was not concerned with a real fluid, but with the hypothetical medium postulated by Descartes in his vortex theory of the universe. Osborne Reynolds, in his last publication, actually worked out a theory of the universe on the assumption that it was a collection of spheres in contact (1903): of the paper in question J. J. Thomson has written (1936), "This is the most obscure of his writings, as at this time his mind was beginning to fail."

The simplest example of dilatancy is offered by a close packed assembly of spheres, in which the arrangement is that of the shot in a "pile of shot". Such an assembly has a bulk density of $\pi/3\sqrt{2}$ if the density of a single sphere is taken as 1. If the overall shape of the assembly is changed in such a way that the spheres adopt the simplest cubic formation, the bulk density becomes $\pi/6$, and there is an increase of volume, or dilatation, of 41%. Dilatation in this sense, which appears to have little to do with the dilatancy of Freundlich, is the subject of this paper. We have already embarked on an attempt to elucidate the dilatancy of Freundlich and hope to be able shortly to present results.

Reynolds exhibited the fundamental feature of dilatancy by filling a rubber bag with spherical shot, the interstices of which were occupied by water. If the shot are close-packed, with no excess water, and the opening of the bag is closed, then the system is rigid in its behaviour, the reason being that the atmospheric pressure resists any rearrangement which would result in an overall increase of volume. If, however, the opening of the bag is connected to a graduated vessel which is open to the air, the bag can be readily deformed. Starting with the bag in a spherical form and with the shot in a close-packed arrangement, Reynolds showed that squeezing the bag between boards resulted in an overall increase of volume which was made evident by the entry of water into the bag; if the deformation were continued, the increase eventually ceased and was followed by a decrease, made evident by water leaving the bag. Before the original volume was reached, this decrease was followed by a second stage of increase. Similar results were obtained with sand. Thus, deformation of a close-packed mass of particles results in a loose-packed mass which can be partly restored to the close-packed condition by further deformation. Reynolds also, as is well known, gave an explanation, based on dilatancy, of the drying which results from the pressure of the foot on water-saturated sand left by the tide.

Jenkin (1931), stimulated by Osborne Reynolds' work, studied the behaviour of a wall of steel discs piled one on the other, with their circular edges in contact, between two retaining uprights, one of which could be moved so as to compress or expand the assembly. He observed the formation of systematic gaps. His interest was in the influence of dilatancy on the pressure exerted by sand masses and the like, and consequently his attention was largely concentrated on the "arching" which takes place in packed masses of particles.

When the present work, which was started early in 1946,* was largely completed, Brown and Hawksley (1947) published an account of experiments performed with a single layer of ¼ in. steel balls arranged on an inclined glass plate. The balls were contained by a fixed horizontal base and two parallel rigid sides, which could be rotated about pivots fixed in the base, so as to deform the assembly. They observed that deformation of packed spheres led to a non-uniform distribution of voids and that areas of tightly packed spheres did not mesh with one another.* They noted lines of slip caused by the movement of large groups of spheres, but found the position of these slip lines to be, in general, indeterminate. The one example in which the slip lines appeared to be determinate corresponded to the arching over a hole in the base, which Jenkin investigated.

---

* The research laboratory at University College was destroyed in 1941 by air attack, and it is only recently that facilities in any way adequate for research have been available there. In 1946 there was a good deal of sand about, but little else, which suggested the theme. The completion of this work has been delayed by prevailing conditions.

In the work to be described we have been particularly concerned with the dilatation produced by the pressure of a piston on the central part of the free horizontal face of an assembly of particles retained by a base and upright walls. The general interest of this case is that the distribution of stresses in the mass, before general disturbance takes place, can be calculated from the theory of elasticity, so that the original type of stress that produces what we may call large irreversible yield, as against elastic yield, can be connected with the nature of the yield. Certain preliminary experiments carried out with sand are first discussed. There follows an account of a precise study of the phenomenon with a two-dimensional hexagonal array of uniform cylindrical rods. A few experiments with three-dimensional irregular arrays are described. The general result is that, in the regular array, deformation proceeds by alternate slipping and healing on three possible planes of easy slip, the alternate slipping and healing being attended by increases and decreases of dilatation. The initiation and course of the slip have been explained on theoretical grounds. On the basis of this work many features of the dilatation of sand and suchlike masses have been explained, as discussed later in this paper.

§ 2. THE DILATANCY OF SAND

Since it seemed possible that hydrodynamic forces might play a part in the dilatancy phenomena observed with ordinary sand, in particular the drying of the sand under pressure of the foot, experiments were carried out to determine whether the rate of loading played any part in the phenomena.

Ordinary "white" beach sand from Lee-on-Solent was washed and sieved, those particles being selected which passed a mesh of 0·21 mm. opening and were retained by a mesh of 0·16 mm. The sand was placed to a depth of 15 cm. in a circular earthenware trough, 30 cm. in diameter, and saturated with water, care being taken that no air-bubbles were enclosed. The surface was carefully levelled and excess surface water removed by a syphon, followed by blotting paper.

The load was applied by a piston P (Figure 1), the brass disc in contact with the sand surface being 6·3 cm. in diameter. It was linked at E to a horizontal lever AB, pivoted to a rigid structure at A and attached at B by a shackle to a hollow cylindrical vessel W, ballasted to float on water in a cylindrical container C. The water could be withdrawn by the tap T.

If the withdrawal of V cm³ of water produces a displacement d of the piston, the point B will move downwards by $D = dL/l$, where $AB = L$, $AE = l$, and the upthrust on the floating cylinder will be reduced by

$$\frac{R_{\mathrm{I}}^2}{R_0^2 - R_{\mathrm{I}}^2}(V - \pi D R_0^2),$$

where R_0 = internal radius of outer can, and R_{I} = external radius of inner can, so that the load on the sand increases up to

$$(V - \pi D R_0^2)\left(\frac{R_{\mathrm{I}}^2}{R_0^2 - R_{\mathrm{I}}^2}\right)\left(\frac{L}{l}\right)$$

---

[*] The senior author has been in the habit for many years of demonstrating this with an assembly of carbon microphone shot between two glass lantern-slide plates.

at the end of the interval t during which the water is removed. The interval t can be varied within wide limits, so that the device provides a simple method of varying the rate of load. The quantity D is, of course, governed by the behaviour of the medium under the piston, and the load will, for a given δV, increase or decrease according as δD is less or greater than $\delta V/\pi R_0^2$.

The actual dimensions were $R_0 = 12{\cdot}3$ cm., $R_I = 11{\cdot}5$ cm., $L = 88{\cdot}9$ cm., $l = 30{\cdot}5$ cm. The displacement of the piston was obtained from readings of a dial gauge, graduated in thousandths of an inch, and the displacement D was, as a check, read by a cathetometer.

As a quantitative estimate of the drying of the sand the reflecting power of the surface for visible light was used. Light from an 8-volt, 6-amp. lamp was projected in a converging beam which illuminated an elliptical patch of the surface, with axes about 7 and 3·8 cm., the centre of the patch being about 8 cm. from the centre of the piston. The intensity of the reflected beam was measured by a selenium photocell, on which a diminished image of the illuminated sand was formed. Light of various colours was tried, as was also the effect of polarization, but the ordinary light from the unscreened lamp was found to be as satisfactory as any of the modifications. The current varied between 45 μa. for dry sand and 70 μa. for wet sand.

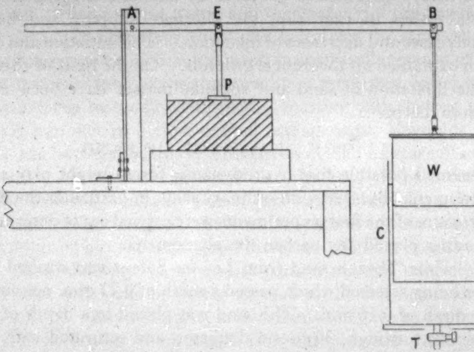


Figure 1. Mechanism for loading the piston P which rests on the sand (shaded).

With this disposition it was shown that the surface dryness produced at a given downward displacement of the piston was independent of the rate of loading over a wide range, a given load, for instance, having been applied in 16 sec. as one extreme and in 600 sec. as the other. It is clear, then, that hydrodynamic forces do not come into the question except, possibly, for rates of loading more rapid than any that we applied.

Before loading, the surface of the sand was level and wet. On loading, the mass of sand was deformed and the surface assumed a new shape, a median section being shown in Figure 2. The central depression represented by the doubly-shaded region in Figure 2 was surrounded by an annular region, shaded singly, in which the sand had been pushed above the initial level. The width of the annular region was about one piston diameter. Outside this region the surface was almost undisturbed. The rise of the sand within the annular region would cause a drying of the surface in this region, irrespective of any dilatation of the sand. Such a local drying would therefore not necessarily indicate dilatation. In the experiments, however, the sand dried over the whole surface, and this extensive drying was taken as evidence of dilatation. The dilatation is probably limited to the neighbourhood of some such region round the piston as is indicated by the dotted curve (Figure 2). In this region the interstices are, on the whole, enlarged, and water flows in to fill them. This redistribution of water causes the whole surface to dry, and the drying in the remoter regions is therefore due to a fall of the water surface below the level of the undisturbed sand.

On withdrawing the piston, the deformation of the sand did not recover completely, a central depression, containing water, remaining in the surface. The whole surface became wet again, but in the remoter surface region it was not quite so wet as in the original state.

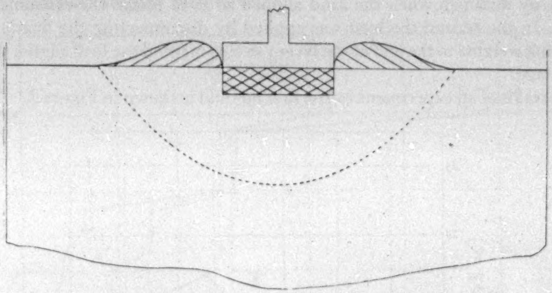


Figure 2. General deformation of sand.

It was found that, after a few cycles of loading and unloading, the sand behaved reversibly, in that renewed application of a given load produced the same drying, and removal of load produced the same wetting. In this reversible condition, the unloaded sand was permanently deformed as compared with its primitive state; its surface was also drier except for the central depression, in which water collected. The given load then produced additional deformation which involved a given dilatation, manifested by a given drying of the surface. On removing the piston, this additional deformation apparently recovered and the surface became as wet as it was before the load was applied. It should be pointed out that, as the piston descended on to the sand, the water which has collected in the central depression was momentarily forced out over the surface of the sand. In consequence the photocell current increased at first, corresponding to this wetting of the surface, which, of course, had nothing to do with dilatancy. As soon as the sand began to be deformed, the surface began to dry, and continued to do so until the full load was applied to the sand.

§ 3. INCREASE AND DECREASE OF DILATATION IN SAND

We have already quoted Osborne Reynolds' experiment demonstrating alternate dilatation and shrinkage of a mass of sand. With the disposition just described, the photocell indicated a progressive drying, the current diminishing until it reached a constant value. This must not be interpreted to mean that there is no alternation of larger and smaller values of the dilatation, since the diminution of volume might not be sufficient to lower the risen surface of the sand to the water level. It would, of course, be possible, just before maximum dilatation was reached, to add sufficient water to reach the surface, in the expectation of seeing it flood the surface as the pressure of the piston was increased, but the regularity of the surface is not sufficient to justify this method.

We show later that the deformation of the two-dimensional array of rods takes place by alternate slipping and healing on certain planes, and that, corresponding to this alternation, the dilatation increases and decreases. To see if evidence could be found of such slipping and healing in sand, load-displacement curves were obtained with sand which had been carefully dried and then packed as closely as possible by gentle shaking. Two methods were adopted. In the first the load was applied by the hydraulic device, which meant that the load would temporarily diminish when the sand slipped so as to lessen the resistance to the piston. In the second the load was applied by disconnecting the floating vessel and adding weights to the end of the lever; in this method the load cannot diminish at any stage.

The result of an experiment by the first method is shown in Figure 3. There is evidence of three slips, at loads of 15·1, 17·1 and 17·4 kg. respectively, the first two followed by a diminution of load as the sand slips under the piston (area 31 cm²). Two different experiments with the second method are recorded in Figure 4, with marked slips at loads of 16·5 and 22·5 kg. respectively. The slips are larger when the load is maintained, as would be expected. There is satisfactory agreement between the pressures at which slip takes place in the two experiments recorded in Figure 4. That the displacements are somewhat different is easily understood, since it is difficult to ensure that the initial packing is the same in the two cases.

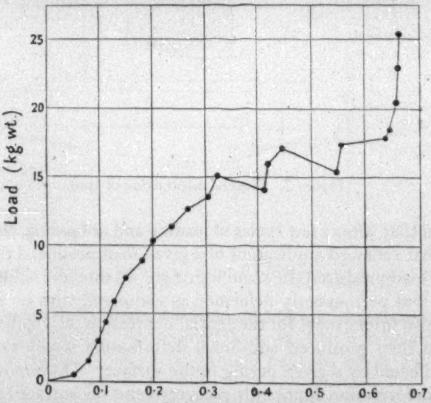


Figure 3. Load-displacement curve for sand with first disposition.

The pronounced hardening denoted by the vertical portions of the curves suggests that the sand under the piston is close packed at these stages.

These experiments indicate clearly alternate slipping and healing in sand, which, no doubt, corresponds to the alternations demonstrated by Reynolds.

§4. THE DILATANCY OF AN IDEALIZED SYSTEM

Sand is too irregular and indefinite a substance to promise precise information on the mechanism of dilatancy. Experiments were therefore undertaken on a regular array of parallel uniform cylindrical rods, which represents a three-fold simplification, viz. uniformity of grain, regularity of packing and limitation of the problem to two dimensions. Subsequently, the behaviour of a two-dimensional array of rods with imperfect packing was investigated. An array of cylinders is better than a single layer of spheres between two glass plates, since there is no tendency for increased friction on the plates when the load is applied. Other advantages readily suggest themselves.

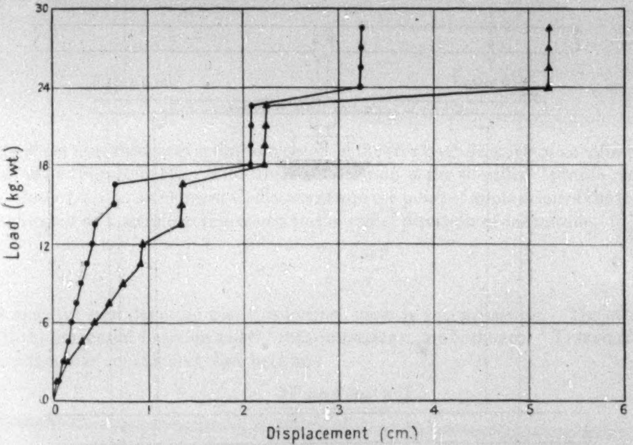


Figure 4. Two load-displacement curves for sand with second disposition.

The experimental arrangement is illustrated in Figure 5. The rods, about 25 mm. long and 4 mm. in diameter, were contained in a Perspex box, of internal dimensions 35 × 35 × 3 cm. Both ebonite rods, which had a fairly rough surface, and polythene rods, which had a smooth, waxy surface, were used. To pack the rods the box was inclined at 45° and the front removed: the rods, some 10,000 in number, were then arranged by hand in 100 layers of 100 rods each and the front clamped on again. For identification purposes every tenth rod in each layer was one whose front was coloured to distinguish it from the other rods, and every tenth layer was composed of these distinguishable rods, so that the undisturbed array showed a square grid, as is seen in, e.g., the parts of Figure 13 and Figure 15 remote from the disturbed area (see footnote, §7). In the case of the ebonite rods the fronts of the marked rods were painted white: in the case of the polythene rods the markers were coloured red with a suitable dye.

The load was applied by a rectangular piston P (Figure 5) constrained to move vertically by a suitable guide. The piston measured about 6·3 by 3·8 cm. and so covered the central fifth of the surface. The piston rod was pivoted to a lever L, and the load applied by the device shown in Figure 1, the water being allowed to escape very slowly. The movement of the piston was stopped at various stages of the disturbance and the position of the rods recorded by photographs, of which Figures 10–13 are typical. This method of applying the load gives perfect control, for when slip takes place the force applied by the piston automatically diminishes, whereas with a dead load, continuously increasing, a step is reached when there is an avalanche, in which the movement of the rods cannot be traced.

With the polythene arrays, front illumination only was used, this being provided by one 100-watt photoflood lamp. The camera was situated at the focus of a single condensing lens of diameter 15 in. and focal length 66 in. placed just in front of the perspex box containing the array. With the ebonite arrays, the reflectivity of the ends of the rods was so poor that front illumination did not distinguish them from the dark interstices. Two methods of illumination proved satisfactory, and both were used. In the first, by means of a pointolite lamp and a large condensing lens, a parallel beam of light was made to fall normally on the back surface of the box. The transmitted beam was focused on the camera lens by a twin condensing lens placed in front of the box. The "grid" was rendered visible by supplementary illumination from the front.

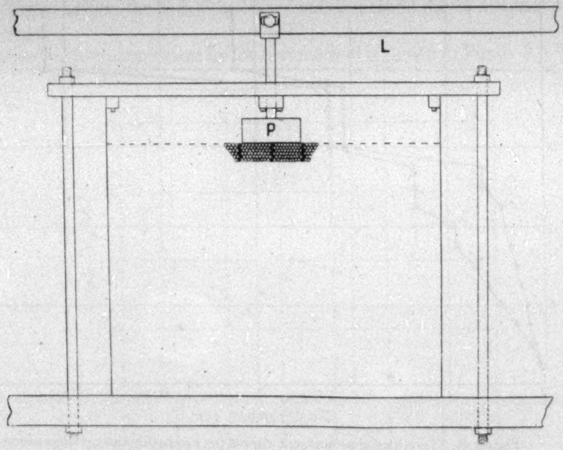


Figure 5. General arrangement for experiments on an array of uniform rods (the array of rods, shown just under the piston, fills the whole box).

In the second method, light from two 500-watt photofloods was scattered between the rods by a special paper (Kodatrace) mounted on the back of the Perspex box. Conditions in front of the box were the same as in the first method.

§ 5. THEORETICAL STRESS DISTRIBUTION FOR ELASTIC DEFORMATION

With the dimensions of the units (rods), box, and breadth of piston in question it is justifiable to apply the theory of elasticity, as worked out for a continuous isotropic medium, to find the distribution of stress in the array before slipping takes place. The problem then is to find the stress distribution in a semi-infinite plate, a portion of whose free edge is uniformly loaded by forces applied normally (Figure 6(*a*)). Use is made of the known solution (Timoshenko 1934) for a vertical force *P* per unit thickness acting at a point on the free boundary AB of an infinitely large plate (Figure 6(*b*)). The stress distribution is the so-called "simple radial distribution", i.e., an element C, distant *r* from the point of application of the load, is subjected to a simple compression in the radial direction of magnitude,

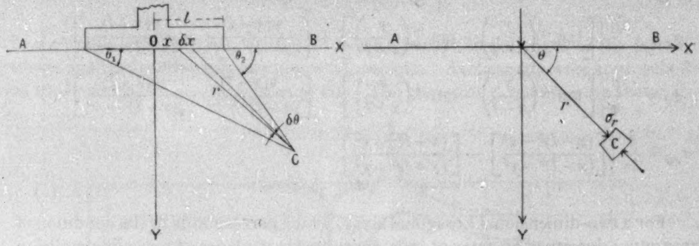


Figure 6 (*a*). Figure 6 (*b*).

For theoretical (elastic) treatment of problem.

$$\sigma_r = -\frac{2P}{\pi}\frac{\sin\theta}{r},$$

the negative sign denoting that this normal stress is a compression. The stress, σ_θ, in the tangential direction and the shearing stress $\tau_{r\theta}$ are both zero. Transferring to rectangular coordinates, this becomes

$$\left.\begin{aligned} \sigma_x &= -\frac{2P}{\pi}\frac{\sin\theta\cos^2\theta}{r}, \\ \sigma_y &= -\frac{2P}{\pi}\frac{\sin^3\theta}{r}, \\ \tau_{xy} &= -\frac{2P}{\pi}\frac{\sin^2\theta\cos\theta}{r}. \end{aligned}\right\} \qquad \ldots\ldots(1)$$

If a uniform load of *p* per unit area is applied by a piston of width 2*l* and unit thickness (Figure 6(*a*)), the load due to an elementary width δx of the piston is $p\delta x = pr\delta\theta/\sin\theta$. From (1) this elementary load contributes a stress component,

$$\delta\sigma_x = -\frac{2p}{\pi}\cos^2\theta\,\delta\theta,$$

on a rectangular element at C, and the whole load gives

$$\sigma_x = -\frac{2p}{\pi}\int_{\theta_1}^{\theta_2}\cos^2\theta\,d\theta;$$

$$\left.\begin{aligned} &\text{therefore} && \sigma_x = -\frac{p}{\pi}[(\theta_2-\theta_1)+\tfrac{1}{2}(\sin 2\theta_2-\sin 2\theta_1)], \\ &\text{similarly} && \sigma_y = -\frac{p}{\pi}[(\theta_2-\theta_1)-\tfrac{1}{2}(\sin 2\theta_2-\sin 2\theta_1)], \\ & && \tau_{xy} = \frac{p}{2\pi}(\cos 2\theta_2-\cos 2\theta_1). \end{aligned}\right\} \qquad \ldots\ldots(2)$$

Expressed in rectangular coordinates, these components are:

$$\left.\begin{aligned} \sigma_x &= -\frac{p}{\pi}\left\{\left[\tan^{-1}\left(\frac{y}{x-l}\right)-\tan^{-1}\left(\frac{y}{x+l}\right)\right]+\left[\frac{y(x-l)}{(x-l)^2+y^2}-\frac{y(x+l)}{(x+l)^2+y^2}\right]\right\}, \\ \sigma_y &= -\frac{p}{\pi}\left\{\left[\tan^{-1}\left(\frac{y}{x-l}\right)-\tan^{-1}\left(\frac{y}{x+l}\right)\right]-\left[\frac{y(x-l)}{(x-l)^2+y^2}-\frac{y(x+l)}{(x+l)^2+y^2}\right]\right\}, \\ \tau_{xy} &= \frac{p}{2\pi}\left\{\left[\frac{(x-l)^2-y^2}{(x-l)^2+y^2}\right]-\left[\frac{(x+l)^2-y^2}{(x+l)^2+y^2}\right]\right\}. \end{aligned}\right\} \qquad \ldots\ldots(3)$$

For a two-dimensional hexagonal array, which corresponds to the condition of maximum density of an array of rods, there are three planes of easy slip parallel to the sides of the unit hexagon. The distribution of shear stress in these planes within the array was worked out on the basis of (3).

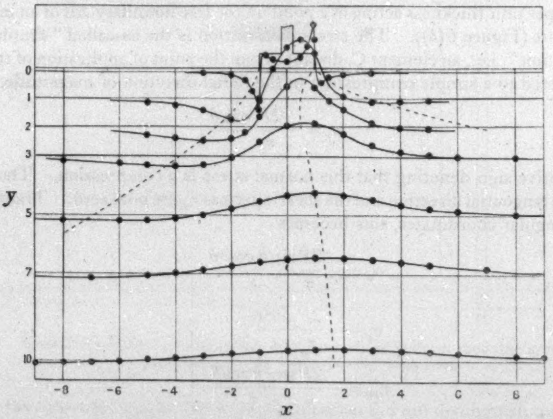


Figure 8. Curves showing distribution of shear stress on planes making an angle of 120° with Ox at various depths below surface.

In the surface plane $y=0$, $\tau_{xy}=0$ everywhere except at the edges of the piston, where there is no unique value for τ_{xy}. On the line $x=l$, i.e. vertically below the right-hand edge of the piston, $\theta_2=\pi/2$ and $\theta_1=\epsilon$, say. From (2)

$$\tau_{xy}=-(p/2\pi)(1+\cos 2\epsilon).$$

Hence $\tau_{xy}\to -p/\pi$ as $\epsilon\to 0$. Similarly, under the left-hand edge of the piston, $\tau_{xy}\to p/\pi$ as $\epsilon\to 0$. For the whole array the maximum value of τ_{xy} is p/π. Curves of τ_{xy} against x are shown in Figure 7 for depths below the surface of 0·1, 1, 2, 3, 5, 7 and 10 units, the piston being taken as 2 units wide. The two symmetrical loci of maximum shearing stress are shown dotted. In these and the remainder of the curves (Figure 8) $p/2\pi$ is taken as the unit of stress.

The two remaining slip-planes in the hexagonal array make angles of 60° and 120° respectively with the x-axis, of which only one need be considered, since everything is symmetrical about the y-axis. The shear stress in the slip-plane making an angle of 120° with Ox is given by,

$$\tau' = \tfrac{1}{2}\tau_{xy} + \frac{\sqrt{3}}{4}(\sigma_y - \sigma_x). \qquad \ldots\ldots(4)$$

In the surface plane $y=0$, $\tau'=0$ everywhere, except at the edges of the piston, where, as already indicated, no unique values exist. As these edges are approached on the lines $x = \pm l$, $\tau' \to \mp p/2\pi$ as $y \to 0$. The curves of τ' against x are shown in Figure 8 and the loci of the three maxima are indicated by dotted lines. These curves are not symmetrical about the y-axis, but the corresponding family for the shear stress in the slip-plane making an angle of 60° with Ox gives rise to another three loci which, with the first three, form a pattern symmetrical about the y-axis.

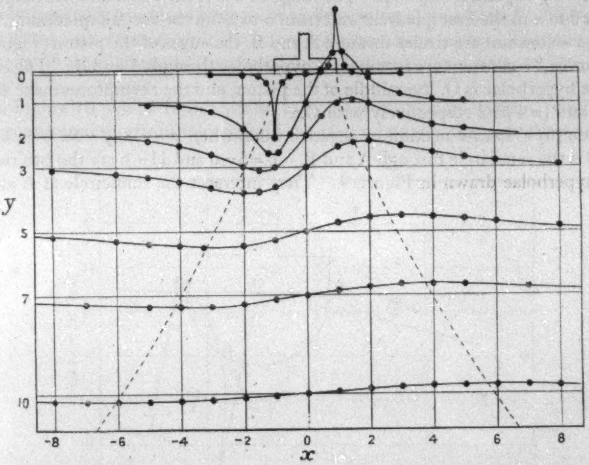


Figure 7. Curves showing distribution of τ_{xy} against x at various depths below surface.

To determine the position of the absolute maximum, τ' is first expressed as a function of θ_1 and θ_2 from (4) and (2), i.e.

$$\tau' = \frac{p}{4\pi}[(\cos 2\theta_2 - \cos 2\theta_1) + \sqrt{3}(\sin 2\theta_2 - \sin 2\theta_1)]. \qquad \ldots\ldots(5)$$

Differentiating with respect to θ_2, keeping θ_1 constant,

$$\frac{\partial \tau'}{\partial \theta_2} = \frac{p}{4\pi}(-2\sin 2\theta_2 + 2\sqrt{3}\cos 2\theta_2).$$

Therefore $\partial\tau'/\partial\theta_2 = 0$ in the directions given by $\tan 2\theta_2 = \sqrt{3}$, i.e., $\theta_2 = 30°$ or $120°$, $\theta_2 > 180°$ having no physical significance.

Similarly, $\partial\tau'/\partial\theta_1=0$ in the directions $\theta_1=30°$ or $120°$. Hence τ' has a stationary value at the intersection of $\theta_1=30°$ and $\theta_2=120°$. At this point both $\partial^2\tau'/\partial\theta_1^2$ and $\partial^2\tau'/\partial\theta_2^2>0$, so that τ' has its maximum negative value at the point. Hence the absolute maximum of the shear stress in a plane making an angle of 120° with the x-axis is at the point $(l/2, \sqrt{3}l/2)$, expressed in rectangular coordinates with origin at O, the middle of the piston (Figure 6 (*a*)). It follows from symmetry that the absolute maximum of the shear stress in the plane making an angle of 60° with the x-axis is at the point $(-l/2, \sqrt{3}l/2)$. The magnitude of the stresses at these points is p/π.

Alternatively* (5) can be written in the form

$$\tau'=-\frac{p}{\pi}\sin\alpha\sin\left(\beta-\frac{\pi}{3}\right), \qquad \ldots\ldots(6)$$

where $\alpha=\theta_2-\theta_1$ and $\beta=\theta_2+\theta_1$. The range in α is from π to 0 and the range in β is from 0 to π in the first quadrant and from π to 2π in the second quadrant. The curves $\alpha=$ constant are circles through A and B, the edges of the piston (Figure 9). The curves $\beta=$ constant are rectangular hyperbolae through A and B. The centre of these hyperbolae is O, the middle of the piston, and the asymptotes make angles of $\beta/2$ and $(\pi+\beta)/2$ respectively with Ox.

From (6) τ' has its maximum negative value when $\alpha=\pi/2$, $\beta=5\pi/6$ or $11\pi/6$. $\alpha=\pi/2$ is the semicircle through A and B. $\beta=5\pi/6$ and $11\pi/6$ are the two rectangular hyperbolae drawn in Figure 9. They intersect the semicircle at A and P.

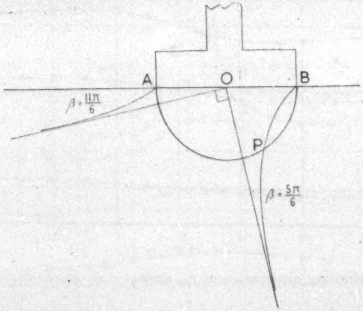


Figure 9. Construction for maximum shear stress on planes making an angle of 120° with Ox.

A lies on all the circles from $\alpha=\pi$ to 0 and on all the rectangular hyperbolae from $\beta=\pi$ to 2π, and so here no unique value of τ' exists. P is the point $(l/2, \sqrt{3}l/2)$, obtained by the previous method.

§ 6. ESTIMATION OF THE DILATATION

In the primitive state of the array, the structure is regular and the piston is just in contact with the level free surface as shown in Figure 5. As the piston penetrates the array, the free surface assumes, at a given stage of the penetration, some such shape as that shown in Figure 2.

The expansion of the array is clearly $(V-V')$ cm^3, where V represents the singly shaded regions at the sides of the piston and V' is the lower part of the piston, shaded doubly (Figure 2). V cannot readily be determined, owing to the irregular upper surfaces of the regions in question. It can, however, be expressed by counting the number R of rods. The volume V', expressed in the same units, is $R' = bdr$, where b is the breadth of the piston, d is the displacement of the piston, and r is the density, in rods per unit area, of the array in the primitive state. The dilatation of the array, expressed in rods, is therefore $(R - R')$. To estimate R at any stage of the deformation, a photograph of the deformed array was used. The edge of a ruler was laid along the level corresponding to the primitive free surface of the array, which was undisturbed in the neighbourhood of the sides of the perspex box, and the number of rods above the ruler was counted, fractions of a rod being estimated to the nearest quarter of a rod.

---

* We are indebted to Mr. A. C. Stevenson for this suggestion.

§7. RESULTS

Figures 10 to 13 (Plate I) show various stages of the deformation of an array of dry polythene rods produced by increasing depression of the piston.* The motion of the piston was arrested while the actual exposures were made. Thirteen photographs in all were taken and measured, of which portions of four are here reproduced.

In Figure 10, which represents an early stage of the deformation, two lines of slip appear, running from the edges of the piston towards the central vertical in a direction making 60° and 120° with the surface. "Holes" caused by misfit have also appeared near the surface to each side of, and beyond, the piston. In Figure 11 a second V-shaped slip system is apparent on the left, and faintly indicated on the right, and horizontal surfaces of misfit have appeared: in Figure 12 almost complete healing has taken place. Photographs not reproduced show stages of renewed slip, each followed by healing. Figure 13 is a stage of advanced penetration, in which many planes of misfit, horizontal and inclined both at 60° and 120°, are visible. The symmetrical nature of the deformation is well shown in this figure.

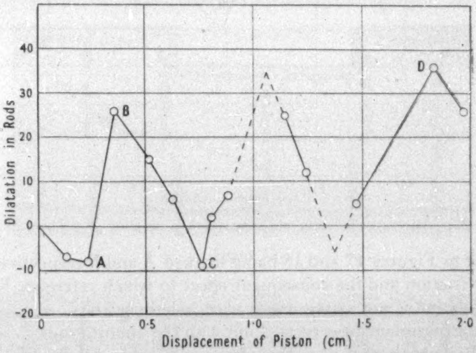


Figure 14. Alternate dilatation and contraction of array of dry polythene rods.

Figure 14 shows the dilatation for increasing displacement in the run from which the photographs here reproduced are selected, the points corresponding to the photographs of Figures 10 to 13 being marked A to D, in order. The alternate nature of the dilatation which corresponds to the formation and healing of the misfit surfaces is clear.

---

* For economy of space, the parts of the photographs showing undisturbed rods to either side and at the bottom have been trimmed off. The actual array was 100 layers of 100 rods each.

Figures 15 and 16 (Plate II) are from another series of experiments on a polythene array. They show very plainly the formation of surfaces of misfit in the directions already specified and the subsequent healing.

A similar mechanism of slip and healing took place with dry ebonite rods, where the friction between rods was much greater than with polythene. Figures 17, 18 and 19 (Plate III) are selected from a large number of photographs, Figure 17 showing many surfaces of misfit, some of which are well healed in the later stage, represented by Figure 18. Reference to Figure 19 is made in the next section, dealing with the mechanism of the deformation. With ebonite, even with the controlled loading adopted, there was a tendency for avalanches to take place, which sometimes resulted in the rods breaking contact with the under surface of the piston. The dilatation for a dry ebonite array is shown in Figure 20, points corresponding to Figures 17 and 18 being marked A and B respectively. Owing to the greater friction and the consequent effect to which reference has just been made, the dilatation is not as regular as with polythene rods.

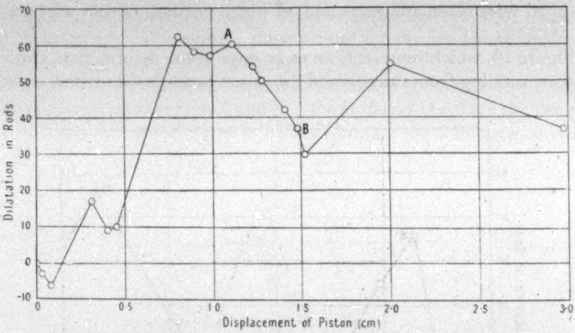


Figure 20. Alternate dilatation and contraction of an array of dry ebonite rods.

A series of experiments was carried out with the ebonite rods in water, which diminished the friction and gave the array some surface cohesion, resulting in greater regularity there. The effects obtained were very similar to those shown by the polythene arrays, and will not be further described. Figure 21 (Plate III) is reproduced because it illustrates particularly well the mechanism of slip now to be discussed.

§8. MECHANISM OF SLIP

In the symmetrical form (which is sometimes realized in experiment, but, owing to chance initial asymmetries, not always) the deformation proceeds in the way indicated in Figure 22. The two equilaterial triangular branches OAB, OA′B′ (Figure 22 (*a*)) of a W-shaped assembly of rods based on the piston move downwards and outwards as units, as indicated by the arrows. These movements are accommodated by adjacent trapezoidal units moving outwards in horizontal directions, that is, ABCD to the right, A′B′C′D′ to the left, and these

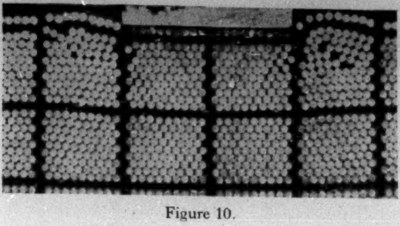


Figure 10.

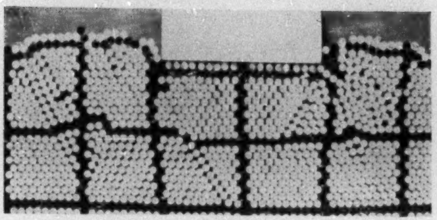


Figure 11.

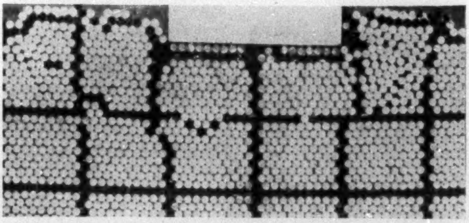


Figure 12.

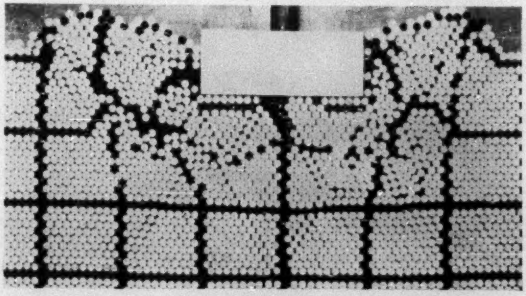


Figure 13.

Stages in the deformation of the array of dry polythene rods in the neighbourhood of the piston.

Plate I.

To face page 496

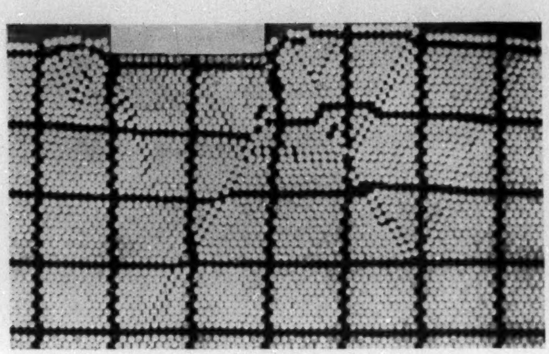


Figure 15.

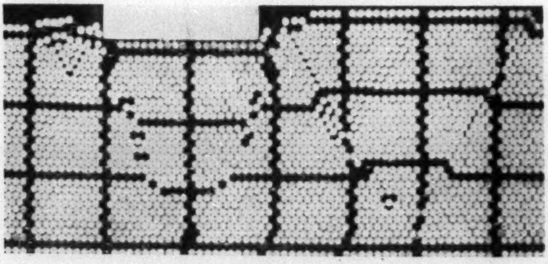


Figure 16.

An array of dry polythene rods during and after slip.

Plate II.

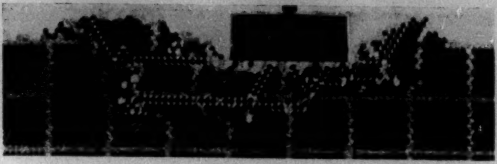


Figure 17.

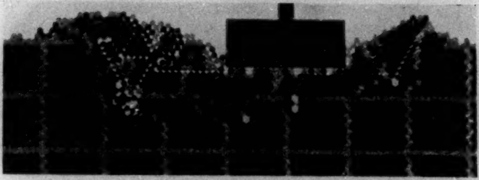


Figure 18.

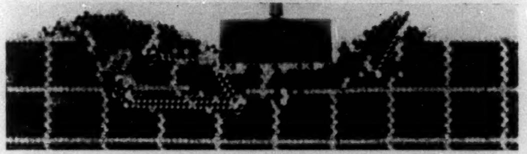


Figure 19.

Deformation of an array of dry ebonite rods.

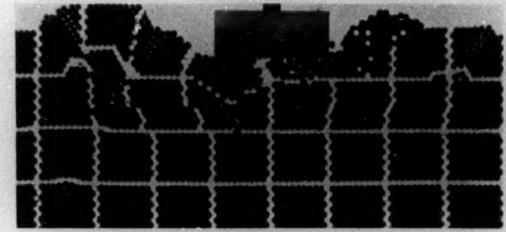


Figure 21.

Deformation of an array of wet ebonite rods.

Plate III.

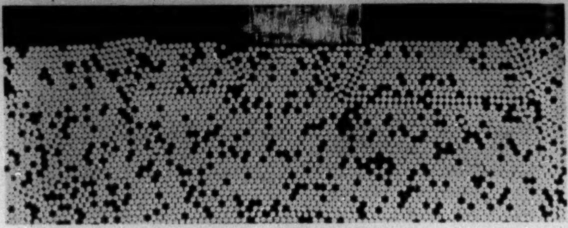


Figure 23.

Deformation of an irregular array of polythene rods.

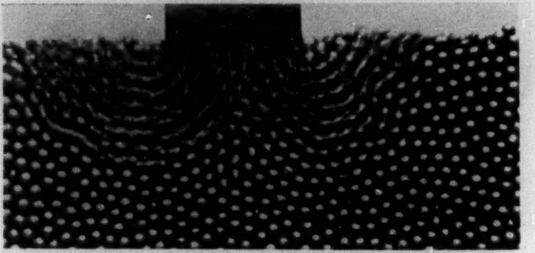


Figure 24.

General movement of carbon shot during descent of piston.

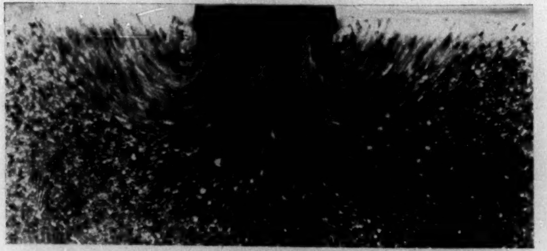
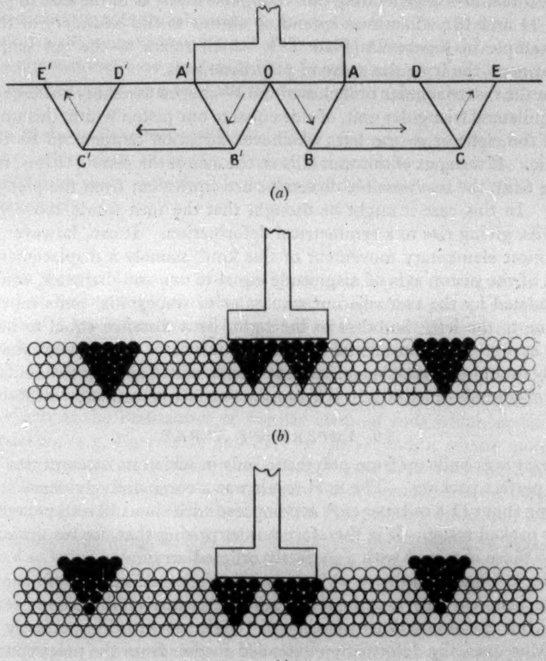


Figure 25.

General movement of sand during descent of piston.

Plate IV.

trapezoidal units push triangular units adjacent to them upwards and outwards, as indicated. The first step, which leaves the rods between the two triangles undisturbed, involves a downward movement of the piston of magnitude $\sqrt{3}r$, where $2r$ is the diameter of the rods. The deformation in terms of rods is indicated in Figure 22(*b*) and (*c*). The dilatation starts at zero, passes through a maximum and, if healing be complete, ends at zero. That there is any dilatation in practice is therefore due to incomplete healing and/or disintegration of one or more of the units involved in the deformation. The lengths AD, A′D′ of the intermediate units ABCD, A′B′C′D′ are variable and may be zero, in which case the trapezia degenerate into triangles.



(*a*)

(*b*)

(*c*)

Figure 22. To illustrate the mechanism of deformation : (*a*) general scheme, (*b*) and (*c*) diagrammatic representation of rods before slip and after complete healing.

This behaviour is explicable in terms of the elastic theory developed earlier. Thus the point B (Figure 22(*a*)) is where the shear stress in the 120° plane has its maximum value. Not only are the 120° and 60° shear stresses a maximum at B and B′ respectively, but these maximum shear stresses occur in the slip-planes BA and B′A′ respectively: the deformation therefore starts by the initiation of slip at B and B′ along BA and B′A′ respectively. Now initiation of slip along BA at B will also initiate slip along BC at B. Once the unit ABCD starts to move by the simultaneous occurrence of slip along BA and BC, the unit OAB will automatically move in the direction shown in Figure 22 (*a*) by slipping along OB. Similarly OA′B′ will slip along AB′.

The deformation in the case of the array of rods will only follow this symmetrical pattern if the array is perfectly homogeneous, a condition which it is difficult to realize. Only in the ideal case will slip be initiated at B and B′ simultaneously. It is not surprising to find that the branches of the W-shaped assembly are often unequal (Figure 21), that slip is sometimes initiated at one point first (Figure 19), and that the triangular branches sometimes coalesce into one larger triangular unit (Figure 15). As already mentioned, the trapezoidal units may vary in width, sometimes degenerating into triangular units, as in the case of polythene (Figures 11 and 15), sometimes extending almost to the boundary of the array, as for example in Figure 23 (Plate IV), which refers to the yet undescribed experiments on the irregular array of polythene rods.

When the two triangular branches of the W-shaped assembly coalesce into one larger, equilateral triangular unit, of side equal to one piston width, this unit moves either to the right or to the left, whichever direction happens to be the more favourable. If the apex of this unit falls on the axis of the piston, that is, the y-axis in Figure 6 (*a*), the two possible directions are equivalent from the point of view of stress. In this case it might be thought that the unit would move vertically downwards, giving rise to a symmetrical deformation. It can, however, be seen that the most elementary movement of this kind, namely a displacement in the direction of the piston axis of magnitude equal to one rod diameter, can only be accommodated by the two adjacent triangular or trapezoidal units moving outwards, one to the left, the other to the right, by a distance equal to half a rod diameter each. These horizontal displacements will, of course, leave the array in an unstable state, and thus the mechanism contemplated appears to be ruled out on the general grounds of stability. This conclusion is in agreement with observation.

§ 9. IMPERFECT ARRAY

An array was built up from polythene rods in which no attempt was made to produce perfect packing. The final result was a comparatively dense structure, there being about 11·8 rods per cm^2, as compared with about 12 rods per cm^2 for the perfectly packed array. It is therefore not surprising that the results were very similar to those obtained with a perfectly ordered array.

Figure 23 (Plate IV) shows a characteristic V-shaped surface of misfit underneath the piston. Alternate increases and decreases of dilatation were again obtained, but the alternations were not as regular as with the perfectly ordered array. Moreover, the deformation extended further from the neighbourhood of the piston.

§ 10. THREE-DIMENSIONAL ARRAYS

The mechanism of deformation in a two-dimensional array of uniform particles having been established, a few experiments were carried out on a three-dimensional array, firstly of uniform spheres and secondly of irregular particles, in order to see if the same general scheme applied.

As spheres, sieved carbon shot were used. They are light, which avoids the complication of large body forces. They were packed in a cubical glass box of side length 5 cm., and the piston was a cylinder, 1 cm. in diameter, half of which had been removed, leaving a plane surface through the axis which was constrained to move vertically, in contact with the glass side through which observations were made. Except for any small changes in frictional effects due to the introduction of the glass in place of a plane of spheres, the conditions observed were therefore those that would be expected to hold in a plane through the axis of a vertical cylindrical piston.

Figure 24 shows a photograph with a time exposure made during the descent of the piston . On this the general movement is indicated by a blurring of the shot in question. An examination of the traces shows that a cone based on the piston, the apex angle of which is 90°, moves downwards; that the adjacent spheres are forced outwards and downwards; while those somewhat further out travel approximately horizontally, as in the two-dimensional case, with the consequence that the spheres more remote travel outwards and upwards. The general picture is a modification in the sense to be expected of the two-dimensional case. To examine the three-dimensional case in detail, both mathematically and experimentally, would have taken more time than we thought well to spend on the matter at this stage.

Figure 25 shows the effect with sand. The process seems to be much the same as with the carbon shot, except for the fact that particles beyond the apex of the right-angled cone show considerable downward displacement. This is probably a frictional effect.

§ 11. DISCUSSION

It is clear that the dilatation of a regular array is intimately connected with the mechanism of deformation, and this mechanism has been demonstrated for the simple case of a two-dimensional hexagonal array when the pressure is applied locally on a free surface, which free surface is parallel to a prismatic face of the hexagons. Two classes of dilatational deformations have been detected, both consequences of slip on well-defined planes. In the first, the dilatation is, as it were, localized at the boundaries of regular areas of rods which preserve their original packing (e.g. Figures 15 and 21); in the second, it occurs more or less irregularly throughout some of these areas (e.g. Figure 13, areas to right and left of piston). The first type is, in a sense, the more fundamental, since the deformation necessarily involves slip which in turn involves this primary dilatation. Whether the secondary dilatation occur or not during the deformation of the array is determined to some extent, at least, by the surface cohesion of the array. If the cohesion be great, then the areas of rods involved in the deformation may be expected to preserve their original packing during slip at their boundaries.

If no secondary dilatation occurred, the primary dilatation would show a fairly regular alternation corresponding to the slipping and healing of the array. The regularity of this alternation is probably disturbed by the secondary dilatation, since this cannot be expected to show any regular variation. On the other hand, the secondary dilatation due to one particular area of rods only must show alternate increases and decreases, since the dilated structure of the area corresponds to an unstable state. Thus the area will expand until it becomes unstable and then collapse into a somewhat close-packed structure.

The experiments on three different regular arrays suggest that the smaller the internal friction, the smaller the rôle played by the secondary dilatation. Thus this type of dilatation was most pronounced in the dry ebonite array, where the internal friction was greatest.

Apparently the internal friction stabilizes the array against the onset of slip, but when the frictional forces are eventually overcome, the deformation becomes catastrophic, with a consequent greater tendency for the areas of rods to break up than when the deformation proceeds gradually, as in the case of small friction.

The wet ebonite array possesses little internal friction and appreciable surface cohesion. Accordingly, the primary dilatation should preponderate, as indeed it did. In the case of the polythene array, the occurrence of the secondary dilatation is to be ascribed not only to the lack of surface cohesion, but also to the slight inevitable curvature of the individual rods. In so far as the latter is involved, it is a spurious effect.

There are certain obvious similarities between the behaviour of the array of rods and that of a single crystal of hexagonal structure, although the inter-particle forces bear little analogy with those that prevail between atoms in a single crystal. Nevertheless, the geometry of the array provides certain planes of easy slip, for which the force required to lift a rod out of the hollow between two other rods and over a boundary into the next hollow is comparatively small. Lines of dislocation roughly correspond to disordered atoms between crystalline grains. Unfortunately, it was not found possible to build with uniform rods a good "polycrystalline" array, which would have provided an analogy to an ordinary metal. It might be possible to simulate the effect of atoms of a foreign solute metal by introducing a small proportion of rods of larger or smaller diameter, which would produce regions of irregularity.

The three-dimensional problem, with irregular packing, is clearly more difficult. To adopt the mathematics of the elastic, three-dimensional case to the problem is a quite straightforward matter but, in view of the time required to work out the experimental case, this has not yet been done.

REFERENCES

BROWN, R. L., and HAWKSLEY, P. G. W., 1947, *Fuel*, **26**, 159.
FREUNDLICH, H., and RÖDER, H. L., 1938, *Trans. Faraday Soc.*, **34**, 308.
JENKIN, C. F., 1931, *Proc. Roy. Soc.* A, **131**, 53.
REYNOLDS, O., 1885, *Collected Scientific Papers*, II, 203; 1886, *Ibid.*, II, 217; 1903, *Ibid.*, III.
THOMSON, J. J., 1936, *Recollections and Reflections* (London : G. Bell & Sons, Ltd.).
TIMOSHENKO, S., 1934, *Theory of Elasticity* (New York and London : McGraw-Hill Book Co., Inc.).

The Effect of Instantaneous Pre-Strain on the Character of Creep in Lead Polycrystals

By A. J. KENNEDY
University College, London

Communicated by E. N. da C. Andrade; MS. received 12*th April* 1949, *read* 4*th February* 1949

ABSTRACT. The extension-against-time curves of lead wires that have been subjected to rapid strain just before the experiment may be expressed by the Andrade creep equation $l=l_0(1+\beta t^{\frac{1}{3}})\exp(\kappa t)$, using the same constants as those which satisfy the creep of the metal under the same constant stress, but with the t value associated with β replaced by $(t+t_0)$, where t_0 is a constant for a given experiment, its value increasing with increasing pre-strain. While for pre-strains less than 10% the value of κ is unchanged by the pre-strain, for greater pre-strains the κ-flow is also modified and appears to approach more nearly to a linear variation with stress.

§1. INTRODUCTION

It is generally accepted to-day that the normal creep of metals comprises two components, a β or transient flow, and a κ or permanent ("viscous") flow (see, e.g. Andrade 1910, Orowan 1947). These two types of flow must correspond to different types of physical change in the metal. It was originally suggested by Andrade (1910, 1914) that the transient flow was connected with a rotation of the axes of the crystal grains, and this hypothesis was confirmed by the work of Andrade and Chalmers (1932) on the resistivity of polycrystalline cadmium, where the fact that a glide process and a rotation of the axis necessarily occur together was clearly brought forward.*

The work of Hanson and Wheeler (1931) and others has made it clear that, in the case of aluminium, deformation by rapid extension at moderate (room) temperatures produces a marked development of glide planes within the crystals, with comparatively little modification in the appearance of the crystal boundaries, whereas slow extension at higher temperatures (250° c.) leads to a broadening and smudging of the boundaries with the ultimate formation of intercrystalline cracks. The formation of a few slip bands was observed in the early stages of the slow extension at 250° c., but they were never comparable in number with those formed at an equal extension at room temperature.

It is, then, well established that, as a first approach, β-flow can be connected with slip within the crystal grains and κ-flow with an intercrystalline movement. At the same time it must be pointed out that glide and consequent rotation of the crystal planes must lead to a modification of the boundaries, and, during the period when the β-flow is taking place, this flow itself affects the intercrystalline region in which the κ-flow is operative. The interaction of the mechanisms is further complicated by the fact that the individual crystals themselves exhibit a permanent, or κ, component, so that boundary flow is not necessarily the only contribution to the polycrystalline κ-flow.

In an earlier investigation (Andrade 1914) it was established that a large rapid preliminary strain, either of extension or torsion, put the metal into a state to flow

---

[*] "The facts can be explained on the assumption that the crystallites slip, with consequent rotation of the unique axis, during the β-flow". Andrade and Chalmers (1932, p. 373). The unique axis was specified because the crystals under consideration were hexagonal.

viscously almost from the start. It was decided to examine this point in greater detail, and to attempt an interpretation of the creep curves in terms of changes in the transient and the permanent components. From the considerations already given it seemed likely that the preliminary strain would produce in the crystals a state closely resembling that normally produced by the β-flow, although it was realized that the distribution of boundary stresses would differ somewhat from that produced by the slower flow, and that the state of break-up of the boundary might also differ. It was hoped that quantitative treatment would throw light upon such points.

§ 2. EXPERIMENTAL METHOD AND RESULTS AT 20° C.

A series of experiments was carried out on polycrystalline lead wires maintained at 19·8° c. in a thermostatic enclosure. The initial lengths of the wires varied from 17 cm. to 20 cm., according to the amount of pre-strain desired, the standard final length after pre-strain being in all cases 20 cm. The axial tension was applied by a constant stress bar (Andrade and Chalmers 1932), the pre-strain being effected by a heavy overload on the bar, which was reduced to that required for the subsequent creep-stress at the instant the specimen attained the standard length. The time taken to effect the pre-strain was in no case greater than 5 sec. Three different stresses, viz. 152·4, 177·8 and 191·6 kg. wt/cm^2, were used for the creep tests and a number of differing pre-strains ranging from 0 to 16%. The load required on the constant stress bar to give the particular creep stress could, of course, be calculated beforehand from the measured initial length of the wire. The lengths were determined at chosen intervals during the flow by a cathetometer reading to 0·05 mm.

The results of the experiments are shown in Figures 1(*a*), (*b*), (*c*), the experimental points being indicated by the circles on the upper curve in Figure 1(*a*) to show the regularity of the readings. It was held to be unnecessary to put in the points in all cases, and for the other experimental curves the individual readings are not shown.

§ 3. EXAMINATION OF RESULTS

A method has been described (Kennedy 1948) whereby the constants in the Andrade creep equation, $l = l_0(1 + \beta t^{\frac{1}{3}}) \exp(\kappa t)$ may be extracted from experimental results without trial and error, so that there is no arbitrary element in the choice of the values of the constants in any particular case. This method was applied to those creep curves for which the pre-strain was zero, and the constants extracted. An attempt was then made to relate the changed character of the creep curves of pre-strained specimens to changes in these values, or to changes in the functions themselves. Examination of the curves provided a simple and consistent explanation of the effect in terms of the creep equation. It was found that all the creep curves, for pre-strains up to 10%, could be closely fitted by assuming the β and κ values to have remained unchanged by the straining, but the β-flow to have been, in a sense, partly exhausted at the start of creep under constant stress, so that the resulting creep curve was the result of a κ-flow determined by the original κ together with a β-flow reckoned from a time preceding the zero time of the experimental creep curve by a calculable amount—an amount which increased with increasing pre-strain*. Thus the values of β and κ for the wire with no

---

* The way in which the fitting was carried out, on this hypothesis, is given in § 6.

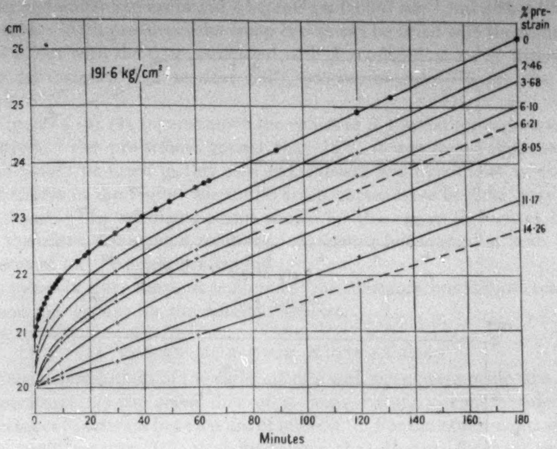


Figure 1 (*a*).

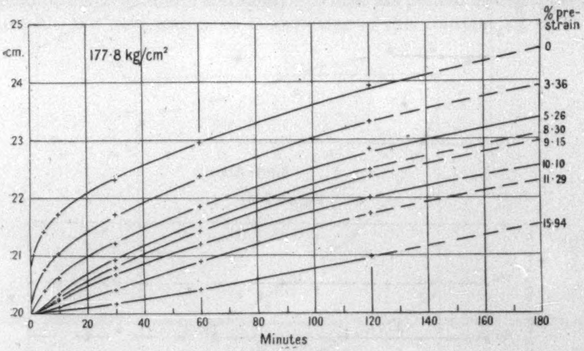


Figure 1 (*b*).

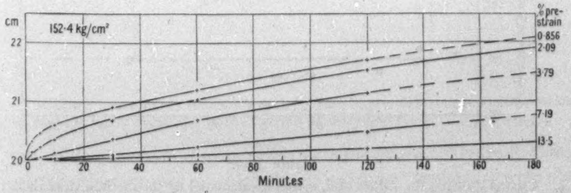


Figure 1 (*c*).

Figure 1. Creep curves of lead at 19·8° C., for the three constant stresses, 191·6, 177·8 and 152·4 kg/cm². The broken lines indicate where the experimental curves have been extended; the crosses indicate the points calculated from the values shown in Figure 2.

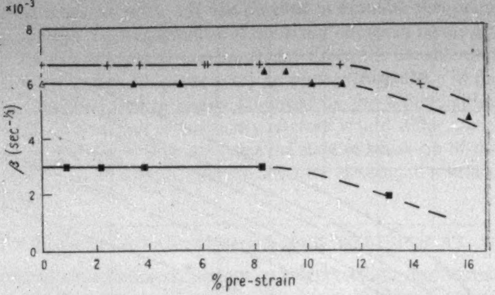


Figure 2 (*a*).

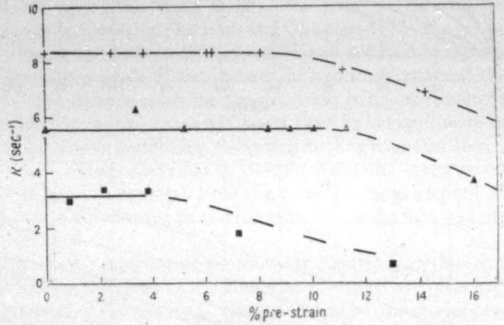


Figure 2 (*b*).

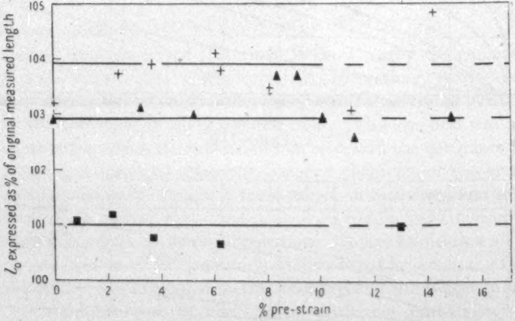


Figure 2 (*c*).

Figure 2. Curves showing the values of β, κ, and l_0 calculated by the method described, for 1 mm. diameter lead at 19·8° c.

+ points given by curves of 191·6 kg/cm² stress.

▲ points given by curves of 177·8 kg/cm² stress.

■ points given by curves of 152·4 kg/cm² stress.

The broken lines indicate the regions within which a unique fitting is not possible because of the straightening of the experimental curves as pre-strain increases.

pre-strain and under a stress of 191·6 kg/cm^2 are 0·0067 sec$^{-\frac{1}{3}}$ and $8{\cdot}36 \times 10^{-6}$ sec^{-1} respectively. With pre-strain the creep curves can be fitted with the same values of β and κ, but with the time associated with β modified to $t+t_0$, instead of t, t_0 being, for instance, 350 sec. for 8·2% pre-strain and 1870 sec. for 11·2% pre-strain.

In Figures 2 (*a*), (*b*), (*c*) are shown the values of β, κ and l_0 for various amounts of pre-strain. For pre-strains greater than 10%, it was found that the curves could no longer be fitted in this way and, indeed, *any* hypothesis is extremely difficult to test in the region where the creep curves have become, very nearly, straight lines. The calculated points shown for this region in Figures 2(*a*), (*b*), (*c*) are, therefore, not unique solutions. It seems, however, that both β and κ must decrease for the larger pre-strains.

The values of l_0 are the most scattered of the constants, but fall into reasonably distinguishable groups for the different stresses.

§4. BEHAVIOUR OF κ WITH STRESS

Various investigations of the variation of κ with stress were made, at a number of temperatures, as the possibility of it being "truly viscous" under certain circumstances has always been a point of interest. For the experiments at 19·8° c. with pre-strain less than 10%, described above, the values of κ for the three stresses used could be extracted quite accurately. These are plotted as the upper curve in Figure 3. We were interested in the value of this constant for an extreme

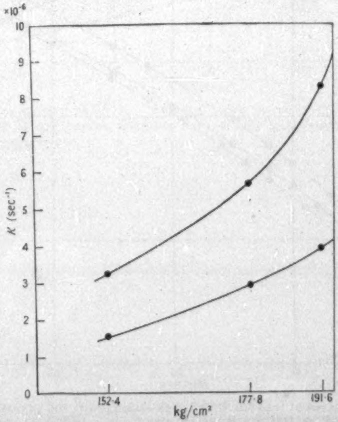


Figure 3. The upper curve shows values of κ derived from Figure 2 (*b*) for pre-strains less than 10%, and the lower curve the values derived as described for 16% pre-strain.

case of, say, 16% pre-strain, but to read this value off from Figure 2 (*b*) would be to overlook the arbitrary way in which these "extreme" constants were obtained. Accepting the validity of the "negative time" concept, a calculation of the κ-values is possible by a simple modification of the fitting method (also described in the note below). The values which this gives are plotted as the lower curve in Figure 3 which shows, by comparison with the curve above it, a marked reduction in curvature. The results of investigations at higher temperatures may be summarized as :

(*a*) In no case can the specimen be pre-strained into such a condition that its rate of flow is proportional to the stress applied.

(*b*) In all cases the values of κ, derived as described, more nearly approach linearity with stress as the pre-strain becomes greater.

(*c*) It is possible that the rates of flow in the *first few minutes* of creep, which, for extreme pre-straining, are less than the steady rate, may repay further analysis on this point.

§ 5. RESULTS OF EXPERIMENTS IN WHICH PRE-STRAINS WERE APPLIED AT TEMPERATURES DIFFERING FROM THAT AT WHICH THE CREEP TEST WAS CARRIED OUT

Following the argument which led to the original experiments, it seemed reasonable to suppose that, provided other complications (such as recrystallization) were not introduced, the effect of pre-strain could be controlled by the temperature at which it was applied, being, it was anticipated, greater for the lower temperatures. Figure 4 shows the effect of applying a 4·45% pre-strain at 100° c., 35° c. and −180° c., and whilst the difference is distinct, the deviation is not great. For a large pre-strain of 9·87% however, as shown in Figure 5, the curves spread out much more. The creep curve of the metal at 35° c., with zero pre-strain, is shown as curve A for comparison. Experiments with liquid nitrogen (−196° c.) gave results indistinguishable from those with liquid oxygen, and there is reason to believe that this curve represents a lower limit reached at temperatures above −180° c.

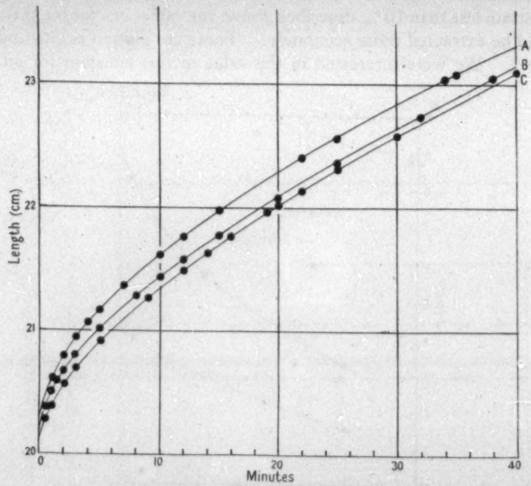


Figure 4. Creep curves of lead for 4·45% pre-strain, with the pre-strains effected for A, at 100° c.; B, at 35° c.; C, at −180° c.

Recrystallization was negligible in these experiments, as the lead used had been heavily contaminated by the addition of tin. The complications which this impurity entailed were accepted, and the consistency of the results seems to confirm that it has been successful in limiting the number of mechanisms that are operative.

§ 6. NOTE ON METHOD OF FITTING THE CREEP CURVES

The method depends simply on the ability to add the β and the κ functions when these are plotted on a logarithmic scale. If we deduce β and κ for the virgin lead by the method already described (Kennedy 1948), we can draw on a diagram similar to that shown (Figure 6) a straight line corresponding to the κ-contribution (OA).

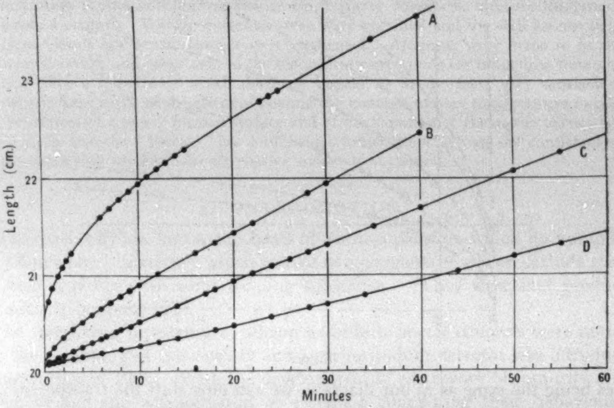


Figure 5. Creep curves of lead at 35° c. A is the creep curve of the normal metal without pre-strain, B, C, and D are curves resulting from a 9·87% pre-strain effected at 100° c., 35° c. and −180° c. respectively.

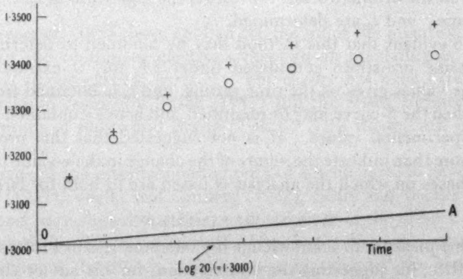


Figure 6

The initial length being 20 cm., the time axis is drawn at $\log 20 = 1{\cdot}3010$. If we now plot on this diagram the logarithms of the measured (experimental) lengths, at various times, of a specimen that has been pre-strained, zero time corresponding to the instant at which it ended its instantaneous pre-strain, we obtain points marked with a cross. Subtracting directly the κ-contribution from each of these points we obtain those marked with a circle.

If we now place over the diagram a transparent sheet on which has been plotted a series of $\log(1+\beta t^{\frac{1}{3}})$ against time curves for different values of β, i.e. a sheet similar to that shown by the dotted lines in Figure 7, both the log and the time scales being the same as in our diagram, we can now shift the transparent sheet about over the diagram until we find the curve which falls along the points marked with a circle. Such a setting gives us at once the data we require, for if this position is that shown in Figure 7, then PB represents the "negative time", and the point B, on the ordinate scale, represents the logarithm of l_0. Thus both the "negative time" and l_0 are determined.

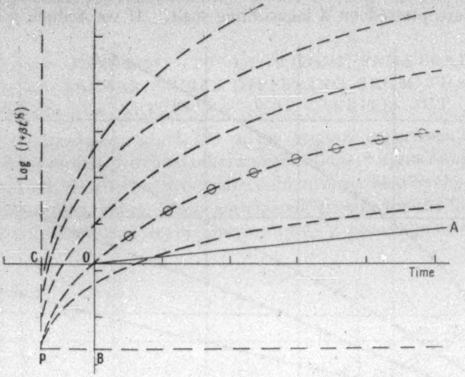


Figure 7

It will be evident that this method may be modified to determine κ for the case of extreme pre-strain considered under §4, for, if extrapolation of the negative time values gives us the time setting, and l_0 is obtained from the Figure 2 (*c*) values, then the β-curve may be identified, and hence κ obtained by subtraction from the experimental values. It is not suggested that this process, and the results, do more than indicate the *nature* of the change in the κ-value that is required if the hypotheses on which the analysis is based are to hold for large pre-strains.

ACKNOWLEDGMENT

It gives me pleasure to acknowledge my indebtedness to Professor E. N. da C. Andrade, F.R.S., for suggesting the investigation, for his advice throughout and for his help with the draft of this paper.

REFERENCES

Andrade, E. N. da C., 1910, *Proc. Roy. Soc.* A, **84,** 1; 1914, *Ibid.*, **90,** 329.
Andrade, E. N. da C., and Chalmers, B., 1932, *Proc. Roy. Soc.* A, **138,** 348.
Hanson, D., and Wheeler, M. A., 1931, *J. Inst. Metals*, **45,** 229.
Hutchison, T. S., 1948, *Nature, Lond.*, **162,** 374.
Kennedy, A. J., 1948, *Proc. Phys, Soc.*, **61,** 510.
Orowan, E., 1947, *J. West Scotland Iron and Steel Institute*, Session 1946–47.

Observations of the Electrical Behaviour of Silicon Carbide Contacts

By E. W. J. MITCHELL and R. W. SILLARS
Metropolitan-Vickers Electrical Co. Ltd.

Communicated by B. G. Churcher; MS. received 18*th February* 1949; *read at Science Meeting at Manchester* 9*th December* 1948

ABSTRACT. The electrical behaviour of individual silicon carbide contacts has been examined using probes and high resistance electrometer circuits to measure the potential drop across a contact. Voltage–current curves were obtained and the well known departures from Ohm's law at the contact were confirmed. Attempts were made to fit these experimental results and other data to the theoretical expressions for tunnelling through an insulating film, a hypothesis which does not appear to fit the facts very satisfactorily. Observations were made of the effect of heating the material to high temperatures *in vacuo*, of the properties of a newly formed surface and of the capacitance across the barrier layer. These indicate that the " barrier " has complicated properties which are not consistent with a simple tunnelling model. An alternative model is suggested.

§ 1. INTRODUCTION

Silicon carbide forms the basis of various resistors which do not follow Ohm's law : materials which consist of aggregates of silicon carbide grains held together with some binding substance. They were first produced commercially in about 1925.

The rectifying properties of silicon carbide to metal contacts were noticed about the beginning of the century and a carborundum detector was introduced by Dunworthy in 1906. With the development of the theory of impurity semiconductors and the accumulation of evidence concerning their behaviour it became apparent that rectification and other contact phenomena exhibited by silicon carbide must be explained on the basis that (*a*) silicon carbide is a semiconductor having a low ohmic resistivity ; (*b*) the surface possesses special properties.

Kurtschatow *et al.* (1935) have applied the theory of the quantum mechanical tunnel effect to the motion of electrons through an insulating layer assumed to exist on each silicon carbide surface. The theory has been extended to the strong-field emission region by Busch (1947) and the result compared with the experimental results of Busch and Braun (1942). This theory will be discussed quantitatively below.

Measurements of the internal conductivity of silicon carbide have been made by Henninger (1937), Voelkl and Guillery (1932), Sears and Becker (1932), and Busch (1946) and have shown that this conductivity is of the order 1 (ohm cm)$^{-1}$ at room temperature. Its variation with temperature has been examined, on a number of specimens, by Busch, who has also investigated the Hall coefficient (Busch and Labhart 1946) and thermo-electric power (Busch, Schmid and Sfondlin 1947) and their variation with temperature. All this work is consistent with the classification of silicon carbide as an impurity semiconductor. Further, the work confirms the conclusion arrived at from earlier observations on thermo-electric power (Hochberg 1937) that black silicon carbide is generally a defect semiconductor (p-type) while the green variety is an excess conductor (n-type). Of the two black varieties commercially available one is characterized by the fact that the "resistance" at a contact is comparatively low ; this will be referred to as "black electrical" grit to distinguish it from the normal "black abrasive". It is used in certain non-ohmic resistors, and is supposed to contain more aluminium impurity than is normally present.

Investigation of the voltage–current characteristics of metal to crystal contacts by Kendall (1944) and Doidge (private communication)—the latter using a probe method—has shown that the observed direction of rectification is opposite to that predicted by the tunnel effect theory applied to such contacts.

The work to be described was nevertheless carried out to test the silica layer–tunnel effect theory in more detail and over as wide a range as possible, since it appeared at one time to be the only way of explaining the behaviour of the contact between two similar semiconductors.

§ 2. VOLTAGE–CURRENT CHARACTERISTIC

(i) *Experimental Results on Single Contacts*

For the single contact measurements the circuit shown in Figure 1 was used. The principal features of this are the electrometer triode bridge circuits which together act as an electrostatic voltmeter placed across the probes. Two electrically homogeneous crystals were used for each determination, one having a large plane face (basal 0001 plane 0·15 cm^2) and the other pointed. The load on the contact was 1·5 kg.

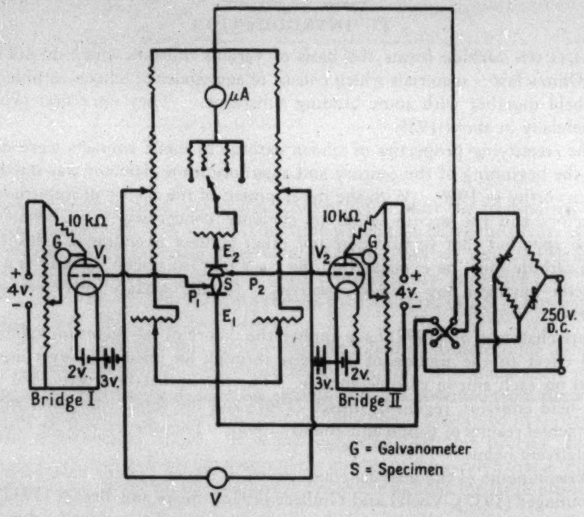


Figure 1. Circuit for measuring volt–current characteristics.

The potential difference between the probes P_1, P_2 was found for currents up to 1 ma. flowing in either direction. The potential drop in the interior of the crystals may be neglected compared with that at the contacts, and one may refer to the potential difference measured between the probes as the voltage drop across the contact. Voltage–current characteristics have been determined for contacts between crystals from the three commercial grades which were available. A typical case is shown in Figures 2 and 3: in Figure 2 the logarithm of the current is plotted against the voltage across the contact for both directions of current, while in Figure 3 the logarithm of the current is plotted against the reciprocal of the voltage.

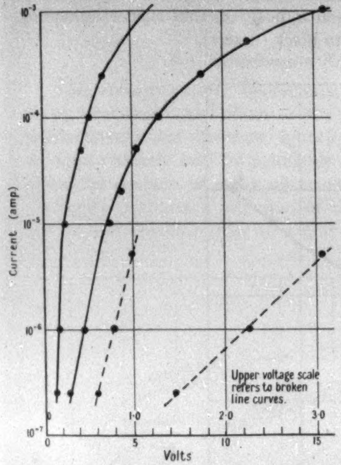


Figure 2. "Abrasive–abrasive" contact.

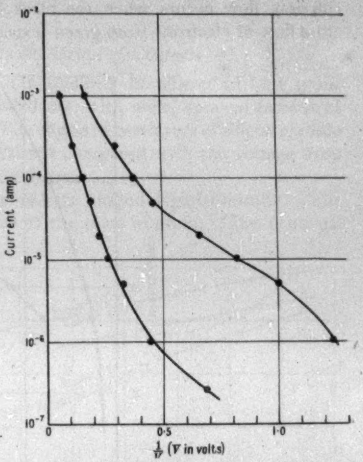


Figure 3. "Abrasive–abrasive" contact.

It is noticeable that in the former the characteristic is approximately linear for potential differences less than about one volt in one direction and three volts in the other, and in this region may be represented by a relation of the form

$$I = Ae^{bV}, \qquad \ldots\ldots(1)$$

where A and b are constants. Values of A, b and of V_c (the voltage below which (1) is approximately valid) for the contacts examined are shown in the Table.

Low Voltage Tunnel Effect

Constants from experimental single contact results.

A – SiC black abrasive,
E – SiC black electrical.

Nature of contact	A–A	A–A	A–A	A–A	E–E	E–E	A–E	A–E
b (volt)$^{-1}$	2·77	2·02	0·65	2·14	1·36	1·50	0·42	1·55
	—	7·42	—	—	2·35	1·57	1·10	1·15
$A \times 10^8$ (amp.)	39	1·7	4·7	0·65	2·65	1·48	30	0·63
	—	0·31	—	—	16	1·30	6·5	3·4
V_c (volt)	1·5	1·0	10	4·0	1·5	4·0	4·0	4·0
	—	3·0	—	—	4·0	4·0	8·0	4·0
$\delta \times 10^7$ (cm.)	1·2	1·1	0·63	1·2	0·82	0·99	0·48	1·0
	—	2·3	—	—	1·1	0·99	0·81	0·84
u (ev.)	1·2	2·0	5·8	2·0	2·7	2·7	7·8	2·8
	—	0·6	—	—	1·5	2·5	3·4	3·4

In the plots of log I against $1/V$ it was apparent that in no region of the depicted experimental range is a relation such as

$$I = A_1 \exp(-b_1/V) \qquad \ldots\ldots(2)$$

valid. A_1 and b_1 are constants.

The characteristic of a green–black contact is shown in Figure 4; the direction of easy flow occurs when the black crystal is positive, that is, corresponding to a flow of electrons from green (excess) to black (defect).

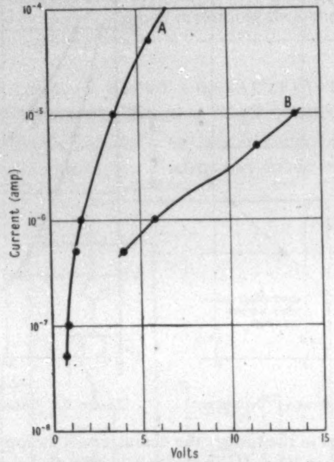


Figure 4. Black–green contact.
Curve A. Black positive. Curve B. Green positive.
Easy direction of electron flow—green to black.

Other points noted from the single contact curves were:

(i) Rectification was always found at "black abrasive" to "black abrasive" contacts, though it is much less pronounced than that occurring at metal to crystal contacts.

(ii) In the case of "black electrical" to "black electrical" this rectification is even less pronounced, and often reverses at higher currents.

(iii) The characteristics of "black abrasive" to "black electrical" contacts lie between the limits of (i) and (ii).

The areas of the contacts were estimated on the assumption that the Hertz relation for elastic compression at contact between a sphere and a plane is applicable. If a sphere of radius R and Young's modulus E rests on a plane surface of the same material under a load P,

$$\text{Area of contact} = 1{\cdot}21\,\pi\,(PR/E)^{2/3} \qquad \ldots\ldots(3)$$

when the area is in cm² if P, R and E are in C.G.S. units.

P was known from the compression of the spring by which the load was applied and E was taken to be 6×10^{12} C.G.S. units. Crystals with "points" which were nearly spherical were selected and measured with the aid of a microscope and photographed at a magnification of 240. The radii of curvature were about 10^{-2} cm.; this leads to areas of contact of order 10^{-5} cm^2 with the loads used (1–2 kg.).

This estimate can be criticized because the compressive stress implicitly assumed is very high. However, no microscopically visible rupture occurred until the load was about 4 kg., and when it did occur the visibly damaged area was about 10^{-5} cm^2 in extent.

(ii) *Experimental Results on Arrays of Contacts*

The voltage–current characteristic of a quantity of silicon carbide grain pressed between metal plates or of a bonded disc with metal sprayed faces is of similar form to that shown in Figure 2. The effects of two layers of silicon carbide to metal contacts may be justifiably neglected compared with the voltage drop across the contacts of many silicon carbide layers.

Figure 5 shows a selection of these results plotted logarithmically. The voltage and currents refer to a unit cube of the mass of grain. The principal features of Figure 5 may be summarized as follows : (*a*) the linear region for voltages corresponding to the order 0·01 volt per contact ; (*b*) the gradual transition between (*a*) and a third region (*c*) ; (*c*) a region extending for many powers of 10 over which a power law,

$$I = kV^{\alpha}, \qquad \ldots\ldots(4)$$

holds, where $3 < \alpha < 6$.

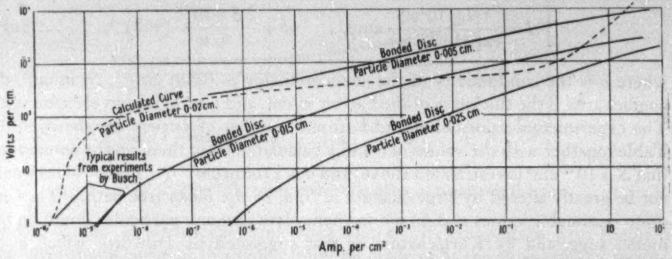


Figure 5. Experimental and theoretical volt-current characteristics of mass of SiC grain.

(iii) *Comparison with Theory*

Theoretical studies of the electrical properties of silicon carbide contacts have been based on the quantum mechanical tunnel effect. They have considered the motion of electrons through a potential barrier arising from an assumed insulating surface layer.

The case where the applied voltage is less than the height of the barrier has been calculated by Kurtschatow *et al.* (1935) and also by Dilworth (1948) ; their results are essentially the same and may be written

$$\text{Current through contact} = I = Ae^{bV}(1 - e^{-40V})\ \text{amp.} \qquad \ldots\ldots(5)$$

at room temperature, where V is the applied potential in volts.

According to this equation there will be a point of inflection in the neighbourhood of 0·1 volt unless b approaches 10 or more. The point of inflection is not found experimentally although the Table indicates that b usually has values of order unity. Busch and Braun (1942) show volt–current curves of aggregates of crystals covering the range 0·001–0·1 volt per contact which are quite ohmic. Kurtschatow *et al.* (1935) fitted their experimental results, also on aggregates of crystals, to a relation $I = \text{const.}\ Ve^{bV}$, which does not show the inflection. Further, Figure 5 also shows the characteristics of a bonded disc of silicon carbide crystals over the range 0·01–0·02 volt per contact, there being no inflection present. It may be that the inflection is masked by the intrinsic conduction of the silica layer itself ; Busch calculates the resistance per contact to be 10^9–10^{11} ohms.

Considering now the region 0·5 volt per contact and above, we take the lower, and roughly rectilinear, parts of Figure 2, assume that the tunnelling mechanism applies, and use the values of A and b of equation (5) to calculate the height and thickness of the barrier layer. The formula of Kurtschatow, with appropriate numerical values inserted, for room temperature, and assuming a mean free path of 10^{-6} cm., gives

$$A = \frac{1{\cdot}9 \times 10^4 \sigma S}{\exp(10^8 \delta \sqrt{u})}\ \text{amp.}\ ; \qquad b = \frac{2{\cdot}5 \times 10^7 \delta}{\sqrt{u}}\ (\text{volt})^{-1}, \quad \ldots\ldots(5a)$$

where σ is the conductivity of the semiconductor in (ohm cm)$^{-1}$, S, in cm^2, the contact area, δ the thickness of the barrier in cm. and u its height in electron volts. The experimental values of A and b from a number of curves are shown in the Table together with the values of δ and u calculated from them on the assumption that S is 10^{-5} cm^2 (as estimated above) and σ is 1 (ohm cm)$^{-1}$. These values would not be greatly altered by large changes in S, σ, or the mean free path. They are quite reasonable values and do not in themselves cast any particular doubt on the model suggested by Kurtschatow or that suggested by Dilworth which gives quantitatively similar results.

It is only when a wider range of the voltage–current characteristic is examined that difficulties arise. Figure 2, typical of many contacts, shows that an exponential relation cannot be fitted except over a very limited range of currents. The same conclusion is illustrated in another way by the dotted curve of Figure 5, which represents the characteristic of a mass of grains, calculated from a tunnelling model by Dilworth, and indicates a far more rapid change of current with voltage than in fact occurs, over a wide range of currents (10^{-6} to 10^{-2} amp. cm^{-2}).

Another feature of this calculated curve of Figure 5 is that between 1 and 10 amp. cm^{-2} the slope changes, becoming 45°—that is, the characteristic of an ohmic resistance. This of course occurs with any model when the "barrier resistance" becomes less than the "constriction resistance" of the contact without the barrier. From the fact that the experimental curves do not behave in this way one must conclude that the area of contact must be larger than was assumed in Dilworth's calculation. This is evidence against the particular geometry of Dilworth's model, but not against the fundamental mechanism of tunnelling. If, however, a larger contact area is assumed, the discrepancy mentioned in the previous paragraph is further emphasized.

The calculations we have been discussing are not valid when the applied potential is greater than the barrier height. The limiting condition, when the field is such that strong field emission occurs, has been considered by Busch and

Braun (1942); by inserting the same numerical values as before into their equation we get

$$I = \frac{1{\cdot}9 \times 10^4 \sigma S}{\exp\{2 \times 10^8 \delta u^{3/2} \phi(1{\cdot}7 \times 10^{-4}(V/\delta u^2)^{1/2})\}/V} \text{ amp.,} \qquad \ldots\ldots(6)$$

where ϕ is the Nordheim function (Nordheim 1928), correcting for the effective lowering of the barrier height u by the image force. If this correction is ignored, equation (6) takes the form of equation (2) which, as we have seen from Figure 3 (a typical case), does not fit the experimental curves. This is brought out more clearly by the upper curve of Figure 6 when the function $I' = 1/\exp(1/V)$ is plotted against V, with logarithmic coordinates. By comparison with Figure 5 it is clear that a curve of this type cannot be made to fit the experimental results. Insertion of the numerical values of equation (6) will merely move the curve bodily without changing its shape. Taking into account the Nordheim correction does not improve matters, as is exemplified by the lower curve of Figure 6. This is plotted for

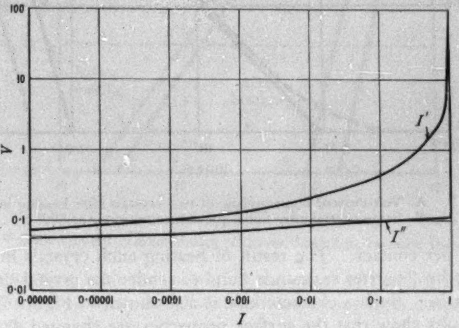


Figure 6. Strong field relation showing effect of Nordheim function.

$$I'' = \frac{1}{\exp\{\phi(3\sqrt{V})/V\}},$$

that is, when $(\delta u^2)^{1/2} = 5 \times 10^{11}$ (e.g. when $u = 2$ ev., $\delta = 6$ A.).

§ 3. EFFECT OF SURFACE TREATMENTS

It was of interest to determine voltage–current characteristics of crystals which had been subjected to treatments such as might be expected to remove the assumed silica layer.

(i) *Heating in Vacuum*

Some crystals were heated in molybdenum boats to a temperature of the order 2,000° c. for about 2 sec. at a time and for about 6 sec. altogether. The pressure in the system—a 70-litre glass bell-jar, pumped by two oil diffusion pumps backed by a mechanical rotary pump—was of the order 10^{-5} mm. Hg as judged from a Philips ionization gauge. The temperature was estimated with the aid of a disappearing filament optical pyrometer. The rate of evaporation of molybdenum from the boat at these temperatures is of the order 10^{-8} gm. per sec.

The voltage–current characteristics of pairs of crystals treated in this way were found and Figure 7 shows a typical case ; this is almost ohmic of resistance

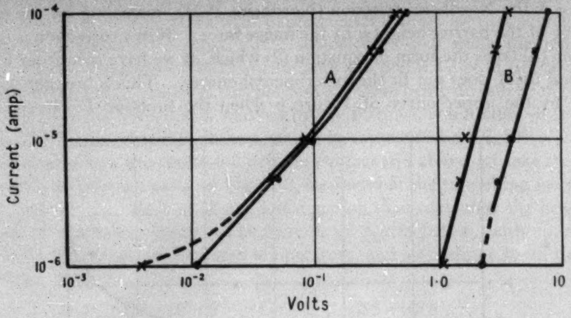


Figure 7. A. Volt-current characteristic of two crystals after heating in vacuum.
B. Same crystal after subsequent heating in air at 800° c.

2×10^4 ohms per contact. The result of heating such crystals in air to 800° c. is to increase the "barrier resistance" and to render the crystal–crystal contacts non-ohmic again. Such a characteristic is also shown in Figure 7.

These results show that the surface properties are changed drastically when the material is heated *in vacuo*. The resistances measured are of the order 100 times greater than the "constriction resistance" calculated from the known conductivity and the estimated area of contact. Thus the process is more complicated than the mere removal of silica. One possibility is that a surface layer of graphite is formed. During the course of the present experiments Busch (1947) has described how the thermal emission plot for silicon carbide which has been heated *in vacuo* to 1,800° c. is identical with that for graphite. The return to non-ohmic conditions after heating ohmic crystals in air to 800° c. is quite consistent with the removal of a graphite layer.

(ii) *Experiments on Freshly Broken Surfaces*

The design of the crystal holder for this work was such that three contacts could be made with the rigidly held crystal. One of them could be removed during the fracturing of the crystal, the other two remaining fixed. The breaking hammer was a rectangular steel bar pivoted at one end about a horizontal axis so that during its descent the protruding crystal was broken.

Voltage–current characteristics were determined for the contacts to the original surface before breaking and to the fresh surface created by the fracture. A metal probe was used as the second member of the contact in one series of experiments and a pointed silicon carbide crystal, 75% of which was immersed in hard solder, in another series. Such changes as were produced in the characteristic by breaking were in no particular direction and were less than the spread, amounting to a factor of 2–3 in the voltage scale, of the characteristics of any metal to crystal or crystal to crystal contact.

Measurements were also made on metal to crystal contacts for crystals heated *in vacuo* as described above. The characteristics of such contacts before breaking were found to be quite ohmic. After fracturing, the characteristic of a contact to the new surface was similar to that of a metal to untreated crystal contact having values of α lying between 3 and 5. An example is shown in Figure 8. This is qualitatively consistent with the view that the effect of heating silicon carbide *in vacuo* is to produce a solid, immobile and conducting layer over the surface—possibly graphite.

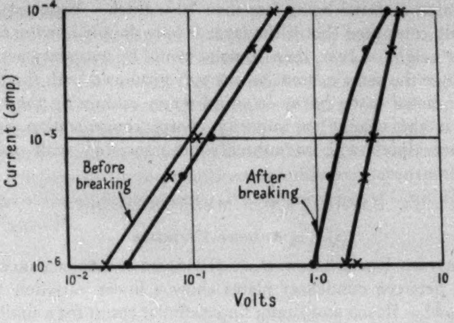


Figure 8. Voltage-current characteristics of crystal heated in vacuum, then broken.

It may be said therefore that, if the properties of the contact are associated with a foreign surface layer, then the layer must be reformed immediately on exposing a fresh silicon carbide surface to air at atmospheric pressure and room temperature.

(iii) *Treatment with Hydrofluoric Acid*

Some authors report that etching with hydrofluoric acid has a profound effect on the electron diffraction pattern of an untreated silicon carbide surface. Claus (1932) claims that the barrier layer can be removed by the same treatment, though the experiments of Kurtschatow do not confirm this. Our experiments with silicon carbide grain pressed between electrodes have shown that there is no significant change between the voltage–current characteristics of grain which has been boiled in hydrofluoric acid solution and those of the untreated material. In fact it is believed that the manufacturers treat at least some of their grades in this manner, yet the electrical properties of all grades examined are essentially similar. It is possible that in some circumstances, if for instance the material had previously been heated in air, etching will produce some modification, but the essential character of the voltage–current curve remains after the most vigorous treatment.

(iv) *Heating in Oxygen*

It is interesting to note that the effect of heating silicon carbide crystals in oxygen to 1,200–1,400° c. for 1–2 hours is generally to increase α in $I=kV^{\alpha}$ and to increase by a factor of 1·5–2·0 the voltage necessary to maintain the original current.

Heine and Scherrer (1940) have found that if silicon carbide is heated in air for a sufficient time, crystallites of crystoballite, identified by electron diffraction and by X-ray powder photographs, are produced. Thus we may consider that one effect of heating silicon carbide in oxygen is to produce a crystoballite layer of increasing thickness. Heine and Scherrer's work suggests that the layer produced by heating to 1,300° C. would be at least several tens of Ångströms thick, any layer on the untreated material being less than 10 A. thick. From equations (5) and (5*a*) it is easily computed that if the layer is only doubled, from, say, 10 to 20 A., with a barrier height of 2 ev., then 10 volts would be necessary across the thicker layer to produce the same current as one volt produced with the thinner.

A further factor to be borne in mind when examining the results of such experiments is the change in impurity centre concentration near the silicon carbide surface which will undoubtedly be associated with oxidation of the surface.

§4. CAPACITANCE MEASUREMENTS

(i) *Capacitance Variation*

Experiments by Kurtschatow *et al.* (1935) on the capacitance of a mass of grain pressed between condenser plates show a linear variation of capacitance with applied load. Busch and Braun find a similar result for a similar experiment at lower loads.

One of us (R. W. S.) has measured the capacitance of bonded silicon carbide discs at low alternating voltages and found it to be independent of frequency. If a D.C. field is superposed however, the capacitance increases with increasing D.C. current and diminishes with increasing frequency. This variation cannot be accounted for on the basis of a rigid silica layer but indicates the presence of a barrier whose thickness, or effective permittivity, or both, vary with the applied field—e.g. some form of space-charge barrier.

(ii) *Single Contact Apparatus*

A method has been developed for the measurement of the small capacitances encountered with such contacts at low voltages, and some preliminary results have been obtained.

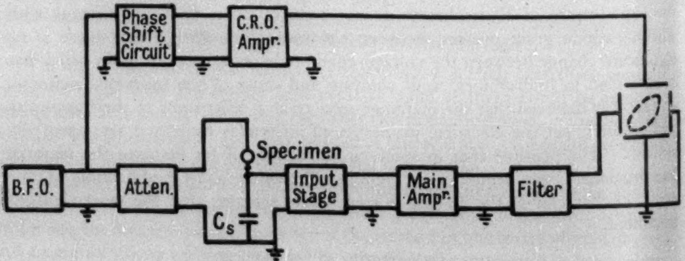


Figure 9. Schematic circuit arrangement for measuring capacitance of contacts.

The apparatus is shown schematically in Figure 9. The crystal was soldered to the central electrode of a cylindrical condenser, C_s, of capacitance 12 pF. The area of the soldered contact was 0·1 to 0·2 cm^2 and it was assumed that the impedance of this could be neglected compared with the impedance of the contact of interest which had an area of only 10^{-5} to 10^{-6} cm. The cathode-ray oscillograph records on one pair of plates an alternating voltage proportional to the impedance of the contact set-up in series with the small capacitance C_s, while on the other pair of plates is a voltage proportional to the impedance of C_s alone. The phase differences between the two voltages are also accurately reproduced on the oscillograph since it was possible to compensate completely at each frequency for phase shifts in the amplifier.

Practically all stray capacitance in parallel with the contact under observation is prevented by the disposition of the earthed metal parts. By leaving the crystals not quite in contact it was verified that the residual strays were negligible.

The traces obtained on the screen were almost but not quite elliptical. In the analysis of the results, mean values of the major and minor semi-axes were used. If we represent the impedance of the contact by an equivalent capacitance C and resistance R in parallel, these quantities are readily computed for the dimensions and orientations of the elliptical traces. The dependence of C and R on frequency is shown in Figure 10.

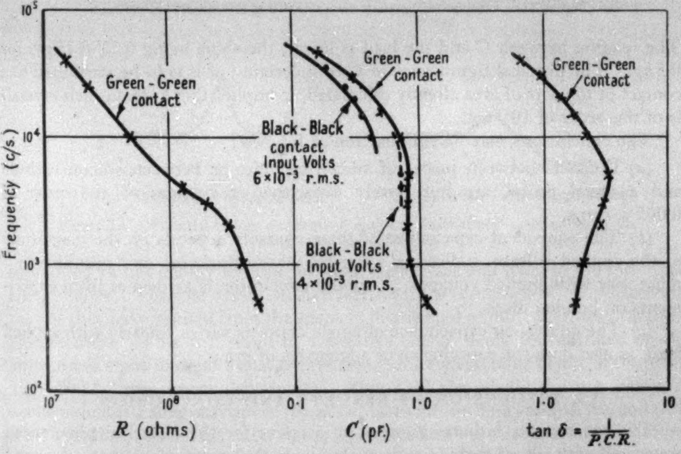


Figure 10. Equivalent resistance, capacitance and loss angle of contacts.

The first point illustrated by these results is that $\tan \delta \simeq 1$. The values of R, ranging from 5×10^6 to 5×10^8 ohms, are obviously not to be identified with the D.C. resistances of the contact, estimated to be approximately 10^9 to 10^{11} ohms at these voltages. Thus the contacts dealt with are not purely capacitive as might have been expected.

Further, the equivalent capacitance C of the contact varies with frequency from 0·1 to 0·2 pF. at 50,000 c/s. to 1 to 2 pF. at 500 c/s., there being a constant region for frequencies between 2,000 and 15,000 c/s. This equivalent capacitance is also dependent on the applied voltage and the mechanical pressure on the contact. The calculated constriction resistance is negligible compared with the observed impedances.

Fragmentary observations showed that increasing the applied voltage results in a decrease in the equivalent capacitance. This change is of the order 50% while the uncertainty in any measurement is estimated to be about ±3%.

The variation of R and C with mechanical pressure is shown in Figure 11.

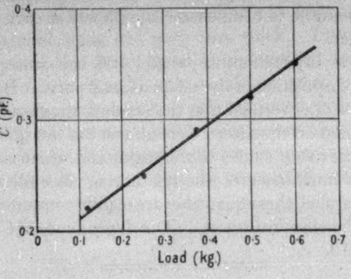


Figure 11. Equivalent capacity versus load for crystal–crystal contact.

The relation between C and the load is linear, the slope being 0·13–0·15 pF. for 0·5 kg. The physical significance of C is uncertain ; if it is to be attributed to a contact of the sort of area already calculated, it implies that δ ($\frac{1}{2}\delta$ on each crystal) is of the order of 10^{-6} cm.

The conclusions may be summarized as follows :

(*a*) Contacts between pieces of silicon carbide, or between silicon carbide and a metal probe, are not purely capacitive at voltages of the order of 0·005 v. r.m.s.

(*b*) The equivalent capacitance of these contacts depends on the magnitude of the applied voltage, indicating a barrier whose thickness or "permittivity" must vary with applied voltage. A similar conclusion is arrived at from experiments on bonded discs.

(*c*) The equivalent capacitance of single contacts varies linearly with applied load, as also does the capacitance of aggregates of grain.

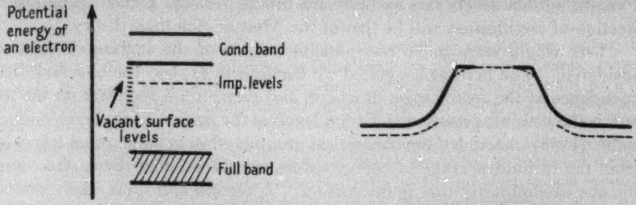
§ 5. DISCUSSION OF RESULTS AND CONCLUSION

The mechanism hitherto adopted to account for the electrical phenomena associated with silicon carbide contacts has been the quantum mechanical tunnel effect, through a potential barrier arising from an insulating layer—presumed to be SiO_2—of thickness about 10 A. assumed to exist on the surface of each crystal.

It has been shown that the observed voltage–current characteristics of crystal to crystal contacts are not satisfactorily explained by this mechanism. Further, the voltage–current characteristics of freshly created surfaces require that any such layer should reform immediately on exposure to air. Treatments which, according to electron diffraction evidence, produce drastic changes in the surface covering, have surprisingly little effect on the electrical characteristics, a circumstance which suggests that this layer is irrelevant. This suggestion is strongly supported by the observations of Jones, Scott and Sillars (1949) that the electrical barrier effects are still present when no surface film more than about 3 A. thick appears to be present.

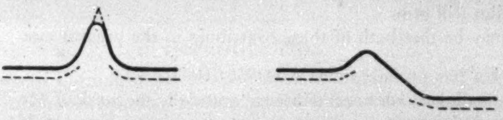
It seems necessary therefore to provide an alternative model for the mechanism of the conduction process across silicon carbide to silicon carbide contacts. The observed changes of capacitance with superposed D.C. field and with changes of applied A.C. voltages suggest that the barrier is not of the rigid type considered above but is rather some form of space-charge barrier.

The manner in which such a barrier may arise has been indicated by Bardeen (1947) and is illustrated in Figure 12. Here (*a*) represents an excess semiconductor containing vacant energy levels at or near the surface and lying in the forbidden energy range. Possible causes of such levels will be discussed below. Electrons may be thermally excited from the sparsely distributed impurity centres into these surface traps. This will result in a negative charge being concentrated on the surface and a positive space charge extending for a considerable depth, of the order 100–1000 A., into the semiconductor. Figure 12 also shows the two pieces of semiconductor a long way apart, (*b*), in contact with no field applied, (*c*), and with a field applied, (*d*). This latter figure illustrates how the thickness and, therefore, the capacitive properties of the barrier may be changed by an alteration in the field applied to the barrier. Thermionic emission will be expected to take place over this barrier as in the Mott (1939) and the Schottky (1941) theories of rectifiers.



(*a*) Excess Semiconductor with vacant surface energy levels.

(*b*) Two such crystals not in contact, showing potential profile of positive charge.



(*c*) In contact – no applied field

(*d*) In contact – with an applied field.

Figure 12. Potential energy diagrams of contact, illustrating the setting-up of the suggested type of barrier.

The problem of explaining the voltage–current characteristic in detail is, unfortunately, not made much easier by this model. The diagram (*d*) of Figure 12 suggests that the characteristics would be similar in both directions to the characteristic of a metal to semiconductor contact in the forward direction. The experimental curves follow an e^{bV} law (ideal forward rectifier characteristic) for only a narrow range of voltage, but it would seem that a more complete theoretical description of the curves must come from modifications of an e^{bV} law rather than from a $(1-e^{-bV})$ law (back characteristic). On the other hand, recent observations on contacts between two similar pieces of germanium (Benzer 1947) indicate that the voltage–current characteristic resembles that of two rectifiers in opposition, which can hardly be explained by the simple barrier of Figure 12 (*d*) ; it may arise in this case from the highly conducting deficit layer which is reported to exist on excess Ge crystals. Thus a contact between two excess crystals is in fact two excess-deficit rectifiers in series opposition.

On the basis of this mechanism the direction of rectification of metal to crystal contacts will be the observed one, for at such a contact will exist a space charge barrier whose height is determined both by the number and energy distribution of vacant surface levels and by the work function of the metal. The resultant direction of rectification will be that of the Mott or Schottky theory.

There would seem to be two possible origins of the surface electron traps considered. The first, as suggested by Bardeen to explain the observed lack of dependence of the rectification of silicon and germanium rectifiers on the metal used, is the possible presence of Tamm levels at the surface of the semiconductor. Tamm (1933) concluded, on theoretical grounds, that in some cases levels lying within the forbidden range of energy values should exist at the surface, arising from the discontinuity that is the surface. A second and more likely source of electron traps is provided by the adsorbed oxygen which must have been present on the surface investigated. In fact the theory of low temperature oxidation of copper and aluminium (Mott 1947), which agrees well with the experimental observations, utilizes such levels to provide a field under the influence of which an oxide film will grow.

It may be that both of these contribute in the present case.

ACKNOWLEDGMENTS

The authors wish to acknowledge gratefully the work of Mr. P. R. Doidge in setting up the D.C. measuring equipment, the assistance of Mr. J. Dunkley in connection witht the capacitance measuring equipment, and the interest and encouragement of Professor N. F. Mott; they also wish to thank Mr. B. G. Churcher, Manager of Research Department, and Sir Arthur P. M. Fleming, Director of Research and Education, Metropolitan-Vickers Electrical Co. Ltd., for permission to publish this paper.

REFERENCES

Bardeen, J., 1947, *Phys. Rev.*, **71**, 717.
Benzer, S., 1947, *Phys. Rev.*, **71**, 141.
Busch, G., 1946, *Helv. Phys. Acta*, **19**, 167; 1947, *Ibid.*, **20**, 33.
Busch, G., and Braun, A., 1942, *Helv. Phys. Acta*, **15**, 571.
Busch, G., and Labhart, H., 1946, *Helv. Phys. Acta*, **19**, 463.
Busch, G., Schmid, P., and Sfondlin, R., 1947, *Helv. Phys. Acta*, **20**, 461.
Claus, B., 1932, *Ann. Phys., Lpz.*, **14**, 644.
Dilworth, C. C., 1948, *Proc. Phys. Soc.*, **60**, 315.
Heine, H. G., and Scherrer, P., 1940, *Helv. Phys. Acta*, **13**, 489.
Henninger, F. P., 1937, *Ann. Phys., Lpz.*, **28**, 245.
Hochberg, B. M., 1937, *J. Exp. Theor. Phys. Sowjet*, **7**, 1090.
Jones, T. K., Scott, R. A., and Sillars, R. W., 1949, *Proc. Phys. Soc.* A, **62**, 333.
Kendall, J. T., 1944, *Proc. Phys. Soc.*, **56**, 123.
Kurtschatow, I. W., Kostina, T. Z., and Rusinov, L. I., 1935, *Phys. Z. Sowjet*, **7**, 129.
Mott, N. F., 1939, *Proc. Roy. Soc.* A, **171**, 27; 1947, *J. Chim. Phys.*, **44**, 172.
Nordheim, L. W., 1928, *Proc. Roy. Soc.* A, **121**, 626.
Schottky, W., 1941, *Z. Phys.*, **118**, 539.
Sears, R. W., and Becker, J. A., 1932, *Phys. Rev.*, **40**, 1055.
Tamm, I., 1933, *Phys. Z. Sowjet*, **1**, 733.
Voelkl, A., and Guillery, P., 1932, *Ann. Phys., Lpz.*, **14**, 193 and 216.

Temperature Measurements of Flames containing Incandescent Particles

By H. G. WOLFHARD and W. G. PARKER
Royal Aircraft Establishment, Farnborough

Communicated by G. E. Bairsto; MS. received 26th January 1959, *read* 11*th March* 1949

ABSTRACT. Colour temperatures measured on photoflash powders are higher than expected from theoretical calculations. Determinations of the true temperature of stationary flames of aluminium flakes suspended in air have been made (*a*) by a line reversal method and (*b*) by measuring the absolute light intensity. The experimental values obtained were very close to the theoretical flame temperature of about 3,000° c., that is, the boiling point of the oxide. An explanation is offered for the difference between the true temperature and colour temperature for flames containing incandescent particles, the reason being the small size of the radiant particles, which introduce optical anomalies. It is concluded therefore to be unprofitable to measure the colour temperature at all as this has no relation to the true temperature. Magnesium behaved very similarly, the true temperature being the melting point of MgO, near 2,800° c., whereas the colour temperature is above 3,900° K. The two-colour method of Hottel and Broughton for temperature measurements in hydrocarbon flames is examined in the light of this experience.

---

§ 1. INTRODUCTION

THE flames of burning metal powders such as aluminium and magnesium exhibit continuous spectra which are probably due to incandescent particles of metal oxide. Determinations of the colour temperature of these flames have been made hitherto on sources of short duration such as the photoflash, and it has been assumed that the particles are Planckian radiators. Caldin (1946) has shown however that if this assumption is justified the values cited for the "temperature" should be increased by 500–1000° K. to obtain the value of the maximum initial temperature. This is because the particles cool down by their own radiation and the spectral distribution of the total energy which is measured is not the same as the spectral distribution corresponding to the initial temperature of the source. The maximum initial temperatures thus obtained are excessively high and above the boiling points of the oxides which are supposed to be the emitters. It is doubtful therefore whether the radiation of these flames can be treated as black-body radiation.

The present authors have also considered this problem and have made some measurements on a stationary flame of aluminium powder burning in air. This flame could be maintained for any convenient time and had a greater optical depth than a photoflash flame. The colour temperature of the flame was found to be considerably greater than the theoretical temperature (which very nearly corresponds with the boiling point of Al_2O_3). Two other possible methods of determining the temperatures were examined, namely the line reversal method and a method based on the absolute light intensity, and the results agreed fairly closely with the theoretical value. The reason for the abnormal colour temperature which is similarly observed in other flames containing small incandescent particles requires an explanation and it has been suggested (Wolfhard and Parker 1948) that the optical properties of very small particles may be responsible.

§ 2. THEORETICAL TEMPERATURE OF THE ALUMINIUM AIR FLAME

The maximum possible flame temperature which can be obtained by a stoichiometric mixture of aluminium powder burning in air (assuming no external heat losses) may be calculated with limited accuracy from the thermal data available for Al_2O_3, viz.

Heat of formation (Rossini and Bichowski 1936) 380 kcal/mol.
Melting point (D'Ans and Lax 1943) 2046° c.
Boiling point (Ruff and Komschak 1926) 2980° c.
Latent heat of fusion (D'Ans and Lax 1943) 26 kcal/mol.
Latent heat of evaporation (Ruff and Komschak 1926) 116 kcal/mol.
Specific heat (D'Ans and Lax 1943):

Temperature (°c.)	−100	0	100	500	1000
Specific heat (cal/gm./°c.)	0·09	0·16	0·22	0·28	0·31

The thermochemical equation which represents the combustion of aluminium in air is

$$2Al + 1{\cdot}5O_2 + 5{\cdot}64\,N_2 = Al_2O_3 + 5{\cdot}64\,N_2 + 380\,\text{kcal}.$$

Figure 1 shows the specific heat of 5·64 mols of nitrogen and 1 mol of Al_2O_3 at different temperatures. The Al_2O_3 value is extrapolated at higher temperatures and it has been assumed that the specific heat of the liquid is similar to that of the solid state. The ordinates of the top curve are the sums of the corresponding ordinates of the N_2 and Al_2O_3 curves. The enthalpy of the system may be computed from the area under the top curve and the heats of fusion and evaporation of Al_2O_3. Figure 2 shows how this enthalpy varies with temperature. Since the heat of combustion of pure aluminium is 380 kcal. the flame temperature just exceeds the boiling point of the Al_2O_3 (2980° c.). When allowance is made for the fact that the metal particles always have an oxide layer (which amounted to 3·9% by weight in the material used for the investigation) it is found that the flame temperature just reaches the boiling point of Al_2O_3 and most of the oxide is initially in the vapour state. At temperatures of this order the dissociation of N_2 is beginning to have a noticeable effect which is observed by the upward bend of the C_p curve for N_2 at about 2500° c. (Figure 1). The influence of the dissociation on the flame temperature is however still small. In this calculation the formation

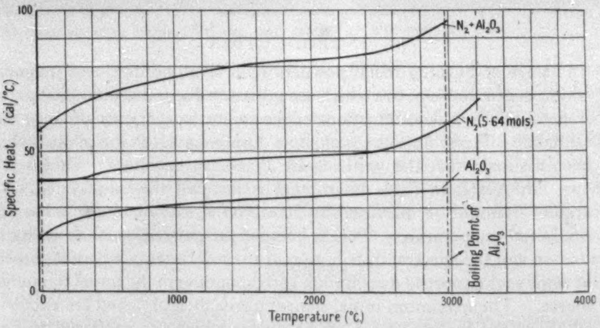


Figure 1. Specific heat (at constant pressure) of 1 mol Al_2O_3 and 5·64 mols N_2.

of nitric oxide is neglected and this is probably justifiable for stoichiometric mixtures. The dissociation of Al_2O_3 is also neglected because of a lack of necessary data. Here again the error is probably small because the Al_2O_3 bands in the spectra of the flame are not particularly strong and such bands are very readily obtained in other light sources containing aluminium. On theoretical grounds therefore we should expect a flame temperature of about 3000° c.

§3. OBSERVATIONS ON A STATIONARY FLAME OF ALUMINIUM POWDER BURNING IN AIR

Aluminium flakes or finely divided metal mixed with air were readily ignited and burned as a stationary flame in a tube. The brilliant white flame could be observed close to the onset of combustion or at any point along its length. The spectrum of the flame principally consisted of a continuum extending from the visible far into the ultra-violet. Al_2O_3 bands and some atom lines of Na and Fe were also present. By the use of a Hilger medium quartz spectrograph and a step sector the continuum was compared with the radiation of the anode of the carbon arc, and an almost constant ratio was found between the carbon arc radiation and the aluminium flame. The colour temperature of the flame was apparently equal to that of the anode of the carbon arc* and amounted to about 3600° c. This is considerably higher than the theoretical temperature even allowing for a reasonable error in the latter.

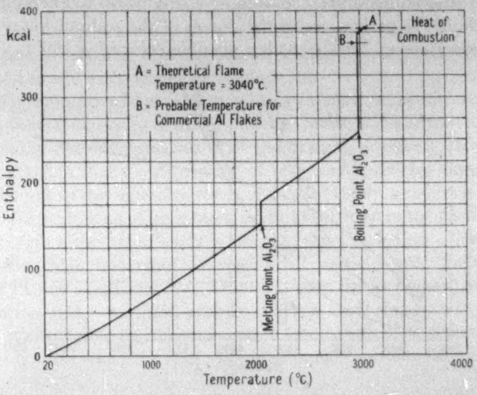


Figure 2. Enthalpy–temperature diagram of the product gases of Al burning in air.

Two possible ways of determining the true temperature were examined experimentally; (*a*) by using a line reversal method and (*b*) by measuring the absolute light intensity.

(*a*) The line reversal method is usually applied to transparent flames. The presence of a continuum in the flame does not necessarily prevent its application because the continuum will reverse at the same temperature as, for example, a sodium line. It is important that the sodium line should be much stronger than the continuum so that the remaining difference effect can be accurately observed

---

* 3818° K. for $\lambda = 6500$ Å. according to Henning and Tingwaldt (1928).

Self-absorption in the cool layer around the flame may be a greater handicap to the method, but this too can be overcome to some extent since the line breadth is greater for the centre of the flame than for the outer layers, and one can observe the reversal desired at the borders of the line.

If the continuum of the flame is emitted by small hot particles there is another source of error in the line reversal method. The theory of the line reversal is based on Kirchhoff's law that the light absorbed is re-emitted. Small particles however scatter the light of the comparative source. There is in the flame therefore, in addition to the true absorption in the particle (which at the reversal point is re-emitted), a pseudo-absorption. The latter is really a reflection and is not re-emitted but scattered. This scattering effectively weakens the intensity of the light source and it is necessary to increase the intensity by changing the sector steps before reaching the apparent reversal point. The measured reversal temperature is probably too high therefore and must be regarded as an upper limit for the true temperature. Assuming a brightness temperature of 3560° C. for the anode of the carbon arc in the region of the sodium line we obtained a reversal temperature of 2900° C. with a possible error of ± 150° C.

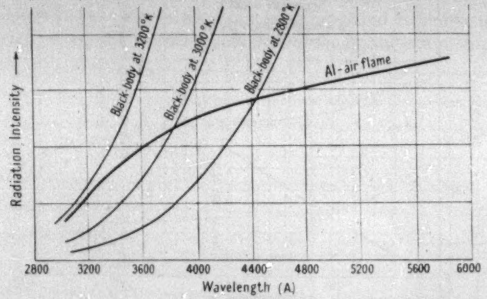


Figure 3.

(*b*) Measurement of the intensity of the continuum in dependence on wavelength is a method of obtaining the temperature of the flame provided the emission or absorption coefficients are known. Determination of the absorption coefficient is difficult however, because the aluminium flame is not sufficiently steady and reproducible. By projecting the image of a carbon arc on to the flame and projecting both on to the slit of the spectrograph it was possible to show that the absorption was small in the visible and increased towards the ultra-violet. At 3000 A. all light from the carbon arc was absorbed and we could therefore estimate a lower limit for the temperature of the flame by assuming that the emission coefficient approached unity at 3000 A. Figure 3 shows the intensity of the aluminium flame plotted against wavelength compared with the intensities of black bodies at 2800, 3000 and 3200° K. respectively. The curve for the flame gradually reaches the values for the 3200° K. radiation, and this temperature therefore provides a lower limit for the true temperature. The intensity curves actually determined varied a little on account of unavoidable fluctuations in the flame. The curve in Figure 3 is the highest which was obtained.

The experimental values found by methods (*a*) and (*b*) which are considered to represent upper and lower limits respectively for the true flame temperature are very close to the theoretical temperature of 2980° C. which is the boiling point of Al_2O_3. We believe that this is the true flame temperature though it differs widely from the colour temperature.

§4. DISCUSSION ON THE FLAME CONTINUUM

If we accept 2980° C. as the true temperature of the flame an explanation is needed of the high colour temperature. In other words, why are the absorption and emission coefficients strongly dependent on the wavelength as illustrated by the divergence of the flame intensity curve from that of the black body at 3200° K. in Figure 3.

The emitter which causes the continuum must be either a gas radiation or hot particles in the flame. Band and line continua have not been recorded which spread over such a broad wavelength region (Finkelnburg 1938) and one is led to assume that the emitters are hot particles. We suggest that these particles are in fact very fine droplets of liquid Al_2O_3 which have condensed from the Al_2O_3 vapour formed in the early stages of combustion.

Oxides of aluminium and similar metals may absorb in the ultra-violet and the emission coefficient naturally tends to increase towards the ultra-violet, but this effect is probably not the explanation of the present problem. It is perhaps a general phenomenon that the absorption of very small particles depends strongly on the wavelength of the light absorbed, and Mie (1908) has shown theoretically that not only the scattering of light but also the absorption in the particles depends on the wavelength if the particle size is less than the wavelength. An explanation of the non-Planckian behaviour of very small particles may lie in Einstein's deduction of Planck's radiation law. Einstein used the Rayleigh–Jeans equations for an evaluation of the transition probabilities and these equations involve a count of the number of possible eigenfrequencies in a body or hollow space. Abnormal conditions may arise when the space becomes equal to or smaller than the wavelength.

Fröhlich (1937) found a similar abnormal effect in the specific heat of very small metal particles when the size of these particles is smaller than the mean free path length of the electrons in the metal (i.e. about 10^{-6} cm.).

Particle sizes in smokes have been determined by a number of investigators and Shechter *et al.* (1946) found from electron micrographs that Al_2O_3 smoke from an arc struck between aluminium electrodes in air consisted of particles of 200–4000 A. diameter. The predominating size was 2000 A. Aggregation may have occurred during the cooling down and sampling of the smoke and the true distribution of sizes may be even smaller. The smoke produced by the stationary flame of burning aluminium powder hung like tobacco smoke in the laboratory and had a bluish colour. Hence we can assume that the particles are smaller than the wavelength of visible light and fulfil, at least partially, the requirements of Mie's formulae. The optical properties of these very fine oxides, particles or droplets depend on the particle size as well as the refractive index and absorption coefficient of the compact material. The overall effect is an absorption which is small in the infra-red but increases in the visible and towards the ultra-violet, and this is what has been observed experimentally. The high colour temperature may be attributed to this abnormal intensity distribution. In bigger flames than those examined in this investigation the optical layer will be greater and the radiation should approach more nearly to black-body conditions on account of self-absorption. The colour temperature will be lower and this indicates how unprofitable it is to measure the colour temperature at all.

§ 5. MAGNESIUM FLAMES

Magnesium behaved in a very similar manner to aluminium. The theoretical flame temperature of magnesium burning in air is approximately 2800° c., which is the boiling point of MgO, and most of the oxide is again formed in the gaseous state. The colour temperature of the radiation of a burning magnesium ribbon appeared to exceed the temperature of the carbon arc comparative source, i.e. above 3900° K. The intensity of the continuum of the flame was compared with black-body radiation of 3000 and 3100° K. (Figure 4). The curve for the flame approached the 3100° K. curve very closely at wavelengths between 2400 and 2800 A., and again we assumed that in this region the emissivity approached unity and the radiation was black-body. The temperature agrees with the theoretical temperature and is close to the boiling point of MgO. The size of the oxide particles from Shechter's measurements with magnesium are in the range 100–10,000 A., and again it is suggested that the colour temperature has no real relation to the true temperature but has its origin in the abnormal emissivity of the tiny oxide particles produced in the flame. The optical depth of the magnesium ribbon flame was less than that of the aluminium–air flame and it is interesting to note that the continuum approached black-body at a smaller wavelength.

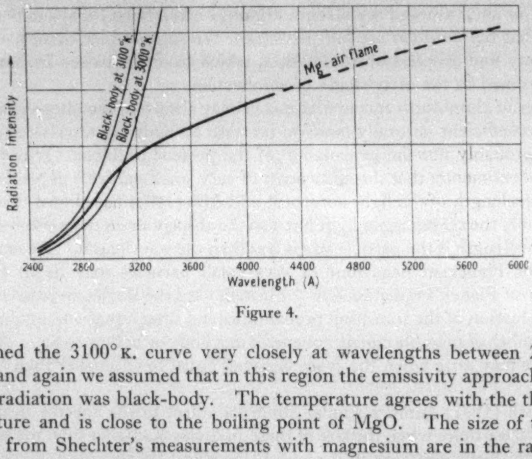


Figure 4.

It is perhaps worth noting that the MgO bands were rather weak in the spectrum of the magnesium flame and another complex band appeared between 3600 and 4000 A. The wavelengths and relative intensities of the more prominent heads of these bands have been listed by Barrow and Crawford (1945) who have also observed the system in absorption. The identity of the carrier is apparently unknown.

§ 6. HYDROCARBON FLAMES

It has long been known that the colour temperature of a Hefner candle is higher than the true temperature and therefore the emissivity of the carbon particles in the flame varies with the wavelength. This is a rather striking fact since bigger pieces of graphite behave as grey bodies. Senftleben and Benedict (1919) who were aware of Mie's theoretical work attributed this difference to the abnormal optical properties of very small particles. This is an important observation for the measurement of temperatures in luminous flames and seems to have been overlooked subsequently. For example Hottel and Broughton (1932) are apparently unaware of it. They measure the brightness at two wavelengths (red and green) and take into account the variation of the emissivity with wavelength in calculating the temperature. They assume that the relationship between the emissivity and the wavelength is simply a property of the material. This is apparently incorrect and the relationship also depends upon the size distribution of the particles, which may vary considerably with the conditions of the flame. When the particles are very small compared with the wavelength, Mie's equations suggest that a single relationship may exist between the absorption of the particle and the wavelength ; and when we have massive carbon the absorption is independent of the wavelength. Between these two extremes the relationship must be strongly dependent on the particle size distribution. Electron micrographs show that the carbon particles in hydrocarbon flames are between 100 and 2000 A. and the sizes are therefore in the critical region. We suggest that the standard two-colour method of measuring the temperature of luminous flames which was introduced by Hottel and Broughton should be examined to see whether these theoretical objections introduce a serious error in the application of the method. It is perhaps significant that the data collected by Hottel and Broughton on measured absorption coefficients shows a remarkable degree of divergence.

ACKNOWLEDGMENT

This paper is based on work carried out in the Ministry of Supply, and the authors are indebted to the Chief Scientist and to the Controller, H.M. Stationery Office, for permission to publish.

REFERENCES

BARROW, R. F., and CRAWFORD, D. V., 1945, *Proc. Phys. Soc.*, **57,** 12.

CALDIN, E. F., 1946, *Proc. Phys. Soc.*, **58,** 341, 350.

D'ANS and LAX, 1943, *Taschenbuch für Chemiker und Physiker* (Berlin: Springer).

FINKELNBURG, W., 1938, *Kontinuierliche Spektren* (Berlin: Springer).

FRÖHLICH, H., 1937, *Physica*, **4,** 406.

HENNING, F., and TINGWALDT, C., 1928, *Z. Phys.*, **48,** 805.

HOTTEL, H. C., and BROUGHTON, F. P., 1932, *Industr. Engng. Chem.*, **4** (Anal. Edit.), 166.

MIE, G., 1908, *Ann. Phys., Lpz.*, **25,** 377.

ROSSINI, F. D., and BICHOWSKI, F. R., 1936, *Thermochemistry of Chemical Substances* (New York: Reinhold).

RUFF, O., and KOMSCHAK, M., 1926, *Z. Elektrochem.*, **32,** 515.

SENFTLEBEN, H., and BENEDICT, E., 1919, *Ann. Phys., Lpz.*, **60,** 297.

SHECHTER, A., ROGINSKY, S., and SAKHAROVA, S., 1946, *Acta Physiochemica, U.R.S.S.*, **21,** 463.

WOLFHARD, H. G., and PARKER, W. G., 1948, *Nature, Lond.*, **162,** 259.

LETTERS TO THE EDITOR

Graphical Representation of a Continuous Noise Spectrum

In plotting an acoustic spectrum in which there is a continuous distribution of energy E with frequency f, the simplest procedure in principle is to plot dE/df (=" the energy per cycle E_c ") as a function of f. However, this is not always convenient, especially where a wide frequency range is involved and E_c tends to a small value at high frequencies as compared with its value at low frequencies. Under these circumstances it is more convenient to plot as a function of $\log f$ a quantity proportional to $dE/d(\log f)$. A parameter which can be used is $f(dE/df)\ln 2$ (=" the energy per octave E_0 ").

This concept of " energy per octave " arises in the following way. For any particular frequency F, consider an octave with frequency limits f_1, f_2 defined by $F^2=f_1f_2, f_2=2f_1$, in which $f(dE/df)$ is taken as constant for the purpose of calculation and equal to its value at F. Then the energy contained in such an octave would be

$$\left(f\frac{dE}{df}\right)_F\int_{f_1}^{f_2}\frac{df}{f}=\left(f\frac{dE}{df}\right)_F\ln 2=(E_0)_F. \qquad \ldots\ldots(1)$$

It should be noted that whereas $E_0/E_c=f\ln 2$, if r.m.s. pressures (P_c, P_0) are used as is often the case in practice, then $P_0/P_c=(f\ln 2)^{\frac{1}{2}}$.

With both the (E_c, f) and the $(E_0, \log f)$ plots, the total energy contained in a frequency region f_A-f_B can be obtained directly by graphical integration. In the case of the $(E_0, \log f)$ plot, the total energy is, from (1),

$$\frac{1}{\ln 2}\int_{f_A}^{f_B}E_0\frac{df}{f}=\frac{1}{\ln 2}\int_{f_A}^{f_B}E_0 d\ln f \qquad \ldots\ldots(2)$$

or

$$\frac{1}{\log 2}\int_{f_A}^{f_B}E_0 d\log f.$$

In the case where we plot E_0 against $\log f$, a frequency interval (e.g. the pass band of a filter) can be expressed as a fraction of an octave. From (2) (with E_0 constant) the frequency interval f_A-f_B is the fraction $\{\log(f_B/f_A)/\log 2\}$ of an octave if we work on an energy basis, or the fraction $\{\log(f_B/f_A)/\log 2\}^{\frac{1}{2}}$ if we work on an r.m.s. pressure basis.

Physics Department,
The University,
Birmingham 15.
3rd June 1949.

D. WALKER.

CONTENTS FOR SECTION A

PAGE
Prof. G. I. FINCH. Steam in the Ring Discharge 465
Mr. M. RYLE. The Significance of the Observation of Intense Radio-Frequency Emission from the Sun 483
Mr. M. RYLE. Evidence for the Stellar Origin of Cosmic Rays . . . 491
Dr. R. R. ROY. Pair-Production in Different Elements by γ-Rays . . . 499
Prof. M. G. EVANS, Mr. J. DE HEER and Dr. J. GERGELY. Structure and Diamagnetic Anisotropy of P-Benzoquinodimethane in Connection with those of P-Benzoquinone 505
Dr. R. MCWEENY and Prof. C. A. COULSON. The Computation of Wave Functions in Momentum Space—I. The Helium Atom 509
Dr. R. MCWEENY. The Computation of Wave Functions in Momentum Space—II. The Hydrogen Molecule Ion 519
Letters to the Editor:
  Mr. K. H. SPRING. Electrons Ejected by Polarized Radiation . . . 529
  Mr. E. SCHWARZ. Theory of Photoconductivity of Layers of Semi-Conducting Substances 530
Reviews of Books 532
Corrigendum 534
Contents for Section B 535
Abstracts for Section B 535

ABSTRACTS FOR SECTION A

Steam in the Ring Discharge, by G. I. FINCH.

ABSTRACT. The behaviour of steam and its decomposition products in the ring discharge has been examined.

Dry hydrogen is not dissociated. The production of atomic hydrogen is dependent upon the presence of steam which dissociates into hydroxyl and atomic hydrogen. A secondary source of atomic hydrogen is then afforded by the interaction of hydroxyl with molecular hydrogen.

The escape from the discharge of atomic hydrogen, a long-lived species, favours the dissociation of steam. Mercury vapour, on the other hand, inhibits the formation of atomic hydrogen and thus leads to a high equilibrium steam concentration.

Unlike dry hydrogen, dry oxygen is dissociated into atoms, but these have a short life as such and recombine in the discharge to form molecular oxygen and ozone.

The reaction mechanisms occurring in the discharge are discussed in the light of spectrographic results.

The Significance of the Observation of Intense Radio-Frequency Emission from the Sun, by M. RYLE.

ABSTRACT. A summarizing account is given of possible mechanisms which have been proposed to account for the emission of intense radio waves from the sun. Whilst most authors agree that the radiation from the undisturbed sun is due to an electron temperature of a million degrees in the solar corona, the greater intensity emitted from sunspots has led some authors to propose non-thermal mechanisms, based on the coherent oscillation of a large number of electrons in the solar corona above the sunspot. The mechanisms proposed for the maintenance of coherent oscillations are examined in detail in this paper and it is concluded that whilst they give adequate explanations for the oscillation observed in discharge tubes, no satisfactory mechanism has yet been suggested for the maintenance of electron oscillations in the solar corona.

It is therefore concluded that the observations can only be accounted for by the occurrence, in the solar corona near sunspots, of electron temperatures of up to 10^{10} deg. K.

The problem of maintaining an electron temperature of 10^6 deg. K. in the normal corona, and 10^{10} deg. K. in the corona near a large sunspot, has been considered in an earlier paper ; it was there shown that by the action of the general magnetic field of the sun and the non-uniform rotation of the solar surface, a potential difference is developed between the poles and the equator which is capable of maintaining an adequate electron temperature in both regions.

Evidence for the Stellar Origin of Cosmic Rays, by M. RYLE.

ABSTRACT. Several authors have suggested the possibility that cosmic rays are due to the acceleration of charged particles in the atmospheres of certain stars. Recent observations at radio frequencies have provided some evidence in support of this hypothesis which is reviewed in this paper.

The observations have shown that at least part of the radio-frequency radiation from the galaxy is emitted by discrete sources of small angular diameter. Measurements of the intensity of the radiation from these sources, and of the time variations of the intensity, indicate that the sources emit radio waves as if they were at a temperature of more than 10^{14} deg. K. It is concluded that this intense radiation cannot be caused by the coherent oscillation of a large number of electrons, but must be due to a genuine electron temperature of about 10^{14} deg. K. (corresponding to a mean electron energy of 10^{10} electron volts). The existence of this electron temperature indicates that in certain stellar bodies there are mechanisms capable of accelerating particles to cosmic ray energies.

A previous theory of the emission of radio waves from the sun is extended to other stellar bodies. It is concluded that a star having a surface magnetic field strength and a peripheral velocity somewhat greater than those which have so far been observed could emit radio waves having the observed intensity. A star of this type could also accelerate charged particles to cosmic ray energies.

The small visual brightness of the sources of radio waves suggests that these bodies may be characterized by small visual opacity or small photospheric temperature.

It is suggested that the difference between the angular distribution of cosmic rays and of the radio waves from the galaxy is due to the deflection of cosmic ray particles by magnetic fields. In addition to previous theories of a general magnetic field in the galaxy, it is shown that the scattering produced by the magnetic fields of the distributed sources themselves should be sufficient to produce an isotropic distribution of cosmic rays at the earth even though the sources may show a marked concentration near the galactic equator.

Pair-Production in Different Elements by γ-Rays, by R. R. ROY.

ABSTRACT. The production of electron pairs on bombardment of Pb, Pt, Sn, Ag and Cu by γ-rays from 200 mg. of Ra and its equilibrium products has been investigated. 1,200 photographs were taken for each element and the yield was 100 pairs in Pb, 85 in Pt, 72 in Sn, 63 in Ag and 50 in Cu. The distribution of energy of the positrons in Pb and Pt follows the shape of the theoretical curve of Bethe and Heitler, but the position of the maximum of the energy distribution curves in Sn, Ag and Cu shows a discrepancy with that to be expected theoretically. The theoretical difference between the average energies of the positrons and electrons in the elements studied, when plotted against Z, is consistently higher than the experimental, though both can be approximately represented by straight lines. The angular distribution of the positrons and electrons with respect to the incident γ-rays agrees with the theory.

Structure and Diamagnetic Anisotropy of P-Benzoquinodimethane in Connection with those of P-Benzoquinone, by M. G. EVANS, J. DE HEER and J. GERGELY.

ABSTRACT. The interest in the diamagnetic properties of quinodimethane lies in the fact that this molecule can be considered as a prototype for quinone. The diamagnetic anistropy, $\Delta\chi$, of p-benzoquinodimethane has been calculated by the molecular orbital method following the treatment given by F. London in 1937. We find :

$$\Delta\chi_{\text{p-benzoquinodimethane}} \sim 0{\cdot}3\Delta\chi_{\text{benzene}}$$

Calculations of the diamagnetic anistropy of quinone itself fail because the perturbation used here breaks down in this case, in which all the centres in the molecule are not identical.

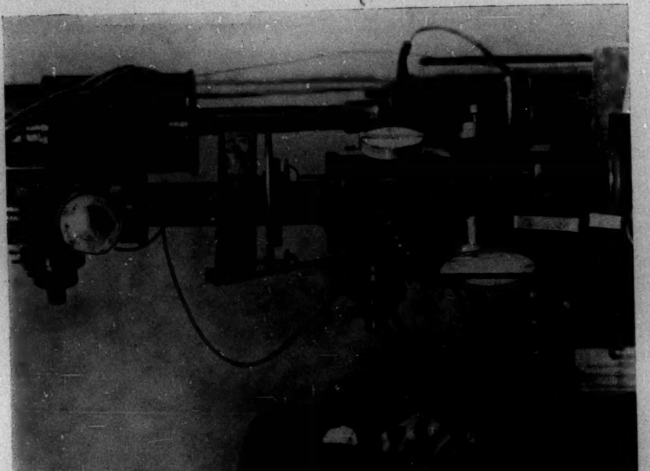
The result obtained for the quinodimethane can be discussed in connection with the observations by K. Lonsdale and K. S. Krishnan on quinone. We conclude that although $\Delta\chi$ of quinone is much reduced compared with that for benzene, we need not postulate, to account for this observation, a complete or nearly complete localization of the π-electrons in this structure.

The Computation of Wave Functions in Momentum Space—I: The Helium Atom, by R. McWEENY and C. A. COULSON.

ABSTRACT. The possibilities of obtaining precise momentum wave functions for atoms by direct solution of the wave equation in momentum space are examined in some detail. An iterative method of approximating to the wave function is employed in computing the momentum distribution function for the helium atom. Although in this instance considerable accuracy is achieved, formidable difficulties arise in extending the calculations to more complicated atoms.

The Computation of Wave Functions in Momentum Space—II: The Hydrogen Molecule Ion, by R. McWEENY.

ABSTRACT. A new approach is made to the problem of computing accurate momentum distribution functions for electrons in molecules. The usual molecular orbital wave function (linear combination of atomic orbitals) has already been employed for this purpose but has proved inadequate in attempts to explain the width of the Compton profile. Using the method of a previous paper, the l.c.a.o. type of molecular orbital is formally corrected by iteration. In most cases it seems likely that the corrected orbital will be associated with a greater spread of the momentum distribution and will thus lead to a broader Compton profile : the broadening should become more marked with increasing delocalization of the electron and may apparently be very considerable. A complete calculation has confirmed these conclusions in the case of the hydrogen molecule ion : a preliminary treatment of the neutral molecule is given.



PROFESSOR G. I. FINCH, M.B.E., D.Sc., F.R.S.,
President of the Physical Society, 1947-1949.